

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 05:50:07 ; Search time 18600 Seconds
(without alignments)
11619.654 Million cell updates/sec

Title: US-09-921-143-36

Perfect score: 5283

Sequence: 1 aagcttgacattatggact.....tgagtgttggcgcagcgtg 5283

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2410	45.6	6074	12	CVE18316	Y18316 Cloning vec
2	2240	42.4	8349	6	AR260588	AR260588 Sequence
3	2233.4	42.3	2661	12	SYN8MRCG	M17626 Synthetic p
4	2218.4	42.0	2999	12	SYNCCDBA	L38498 Cloning vec
5	2218.4	42.0	2999	12	SYNCCDBB	L38499 Cloning vec
6	2177.4	41.2	2604	12	AY189826	AY189826 His-3 int
7	2177.4	41.2	11373	12	AY189827	AY189827 His-3 int
8	2177.4	41.2	11403	12	AY189829	AY189829 His-3 int
9	2088.4	39.5	3036	12	AY222815	AY222815 Cloning v
10	2088.4	39.5	3210	12	AY219859	AY219859 Cloning v
11	2088.4	39.5	5695	12	AY222822	AY222822 Shuttle v
12	2088.4	39.5	5869	12	AY219861	AY219861 Shuttle v
13	2088	39.5	5595	12	AF445080	AF445080 Cloning v
14	2078	39.3	3035	12	AY222814	AY222814 Cloning v
15	2078	39.3	3035	12	AY222811	AY222811 Cloning v
16	1892	35.8	7020	12	AS293724	Z93724 Murine retr
17	1793.4	33.9	4800	6	AR282049	AR282049 Sequence
18	1730.2	32.8	7414	12	AF402285	AF402285 PK[BIG-a]
19	1692.4	32.0	5225	6	AX226280	AX226280 Sequence
20	1671.2	31.6	5041	12	PKSM713	U04895 Cloning vec
21	1671.2	31.6	5041	12	PKSM715	U04896 Cloning vec
22	1671	31.6	4359	12	AY219686	AY219686 Expressio
23	1671	31.6	4968	12	AY219687	AY219687 Expressio
24	1669.6	31.6	7018	12	AY219683	AY219683 Shuttle e
25	1652.4	31.3	5594	6	AR214682	AR214682 Sequence
26	1644.6	31.1	2959	12	AF402779	AF402779 Expressio
27	1625.6	30.8	3741	12	PKSM711	AR24894 Cloning vec
28	1618.2	30.6	6561	6	AR214683	AR214683 Sequence
29	1603.8	30.4	3599	6	BD063790	BD063790 Insulin-1
30	1603.8	30.4	3599	6	BD069041	BD069041 Treatment
31	1602.2	30.3	3600	6	BD063789	BD063789 Insulin-1
32	1602.2	30.3	3600	6	BD069040	BD069040 Treatment
33	1602.2	30.3	3600	6	BD136050	BD136050 Interfero
34	1600	30.3	5966	6	AF098284	AF098284 Cloning v
35	1572.6	29.8	8363	12	AR219825	AR219825 Sequence
36	1571.6	29.7	4665	6	BD082038	BD082038 Immunogen
37	1571.6	29.7	4665	6	BD139247	BD139247 Method of
38	1570.8	29.7	5063	6	BD139247	BD139247 Signal pept
39	1570	29.7	4517	6	E59418	E59418 Sequence
40	1570	29.7	4518	6	AR214679	AR214679 Sequence
41	1570	29.7	4518	12	XXU37573	U37573 Shuttle exp
42	1570	29.7	4886	6	AR214349	AR214349 Sequence
43	1570	29.7	5619	6	AX008865	AX008865 Sequence
44	1570	29.7	5628	6	AX008867	AX008867 Sequence
45	1570	29.7	5775	6	AX641960	AX641960 Sequence

ALIGNMENTS

Result 1	Score	Query Match	Length	DB	ID	Description
CVE18316	2410	45.6	6074	12	CVE18316	Y18316 Cloning vec
LOCUS						Cloning vector pKGM.
DEFINITION						Cloning vector pKGM.
ACCESSION						Y18316
VERSION						Y18316.1
KEYWORDS						bleomycin resistance; cloning vector; kanamycin resistance; multiple cloning site; neomycin resistance.
SOURCE						Cloning vector pKGM
ORGANISM						Cloning vector pKGM
REFERENCE						artificial sequences; vectors.
AUTHORS						1
TITLE						Bannasch, D. and Schwab, M.
JOURNAL						A versatile bait vector for rapid Gal4 dependent two-hybrid screens
						Unpublished

Pred. No. is the number of results predicted by chance to have a

```

REFERENCE 2 (bases 1 to 6074)
AUTHORS Bannasch,D.
TITLE Direct Submission
JOURNAL Submitted (30-OCT-1998) D. Bannasch, Institution Division of,
Cyto genetics/H0400 Deutsches, Krebsforschungszentrum (DKFZ), Im
Neuenheimer Feld 280, D-69120 Heidelberg, FRG
REMARK Revised by author 10-JUN-1999
COMMENT Related sequences U07646, U00004, L19385.
FEATURES
    Location/Qualifiers
        1..6074
            /organism="Cloning vector pKGLM"
            /mol_type="genomic DNA"
            /db_xref="taxon:91498"
            434..967
                /note="ORF1"
            /codon_start=1
            /product="hypothetical protein"
            /protein_id="CAB54527.1"
            /db_xref="GI:5830188"
            /translation="MKGLSSIEQACDIORLKKKCKEKPCKAKLKNWECRYSPT
            KRSLPLRAHLTEVERSLERLEQLFLIPREDLMLKMDSLQDIKALLTGLFVQNV
            NKDAVTRLASVEVDMPILTRQHRISATSSSESSNKGRQLTVSPFPGVRYRRP
            AAKLPGBFMIYDFY"
            434..874
                /note="Gal4 DNA binding domain"
            877..905
                /note="multiple cloning site"
                complement(3126..3506)
            /codon_start=1
            /product="bleomycin resistance protein"
            /protein_id="CAB54529.1"
            /db_xref="GI:5830190"
            /translation="MTQATNPLSRPFDSTAVERLGFIVFRDAGMWILQRGLM
            LEFAHPGLDPLAFWFSQCLRLDLAEFYRQCKSVGIQETSSGYPRIHAPELQWGT
            MAALVDPDGTLLRIQNELLAGIS"
            complement(3527..4321)
            /codon_start=1
            /product="neomycin/kanamycin resistance protein"
            /protein_id="CAB54528.1"
            /db_xref="GI:5830189"
            /translation="MIFQDGLHAGSPAANVERLFYDWAQQTIGCSDAVFLRLSAQR
            FLVFKVDLSGALNELQDEARLNLATTPVCAALVDVVTAGRDWLLIGVPGQDL
            LSHLAPAEKVSIMADARLRLTDPATCPDHOAKHRIERARTMEAGLVDDDDIDE
            BHQGLAPAEFLARLKPMPGDEDLVTHGDACLNINVENGRPSGFIIDCGRLVADRY
            ODIALPEDIARELGGEWADFLVLYGIAPDSORIAFYRLIDFF"
            complement(4410..4459)
            /gene="aminoglycoside phosphotransferase"
            complement(4410..4459)
            /gene="aminoglycoside phosphotransferase"
BASE COUNT 1593 a 1549 c 1380 g 1552 t
ORIGIN
Query Match 45.6%; Score 2410; DB 12; Length 6074;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 2568; Conservative 0; Mismatches 5; Indels 145; Gaps 1;
QY 2711 CGTAATCATGTGTCATAGCTGTTTCTGTGTCAATGTTATTCGCTCACAATCCACACA 2770
DB 1954 CGTAATCATGTGTCATAGCTGTTTCTGTGTCAATGTTATTCGCTCACAATCCACACA 2013
QY 2771 ACATACGACCGGGAAGCATAAAGTGTAAAGCTGGGTGCTTAATGATGAGCTAACTCA 2830
DB 2014 ACATACGACCGGGAAGCATAAAGTGTAAAGCTGGGTGCTTAATGATGAGCTAACTCA 2073
QY 2831 CATTAAATTCGTTGGCTTCATGCGCTTCCAGTCGGGAACCTCTGTCGCCAGCTGC 2890
DB 2074 CATTAAATTCGTTGGCTTCATGCGCTTCCAGTCGGGAACCTCTGTCGCCAGCTGC 2133
QY 2891 ATTAAATGAATCGGCAACCGCGGGAGAGCGGTTTCGCTATTGGCGCTCTTCGCGCTT 2950
DB 2134 ATTAAATGAATCGGCAACCGCGGGAGAGCGGTTTCGCTATTGGCGCTCTTCGCGCTT 2193
QY 2951 CCTCGCTCACTGACTCGCTCGGCTCGGTCGTTGCGCTCGCGGAGCGGTATCAGTCACT 3010

```



```
Db 3274 GATTTCACATGCGGTAGAACTCCGCGAGGTGCTCCAGCCTCAGCGACGACTGAACAA 3333
Qy 3946 CTCGCGAGGGATCGAGCCCGGGTGGCGAAGAACTCCAGCATGAGATCCCGCTGG 4005
Db 3334 CTCGCGAGGGGATCGAGCCCGGGTGGCGAAGAACTCCAGCATGAGATCCCGCTGG 3393
Qy 4006 AGGATCATCCAGCCGCGTCCCGGAAACGATTCGGAAGCCCAACCTTTCATAGAGCG 4065
Db 3394 AGGATCATCCAGCCGCGTCCCGGAAACGATTCGGAAGCCCAACCTTTCATAGAGCG 3453
Qy 4066 GCGGTGGAATCGAAATCTCGTATGCGCAGGTGGCGTGGCTGGTGGTGGTGGTGGTGG 4125
Db 3454 GCGGTGGAATCGAAATCTCGTATGCGCAGGTGGCGTGGCTGGTGGTGGTGGTGGTGG 3513
Qy 4126 CCCAGAGTCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGCGGATGGCTGGCAAT 4185
Db 3514 CCCAGAGTCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGCGGATGGCTGGCAAT 3573
Qy 4186 CGGAGCGCGATACCGTAAAGCAGAGAGCGGTCAAGCCATTCGCGCGCAAGCTCTT 4245
Db 3574 CGGAGCGCGATACCGTAAAGCAGAGAGCGGTCAAGCCATTCGCGCGCAAGCTCTT 3633
Qy 4246 CAGCAATATCAGCGGTAGCAACGCTATGCTCTGATAGCGGTCCGCCACACCCAGCGGC 4305
Db 3634 CAGCAATATCAGCGGTAGCAACGCTATGCTCTGATAGCGGTCCGCCACACCCAGCGGC 3693
Qy 4306 CACAGTCGATGAATCCAGAAAGCGCCATTTTCCACCATCATATTCGGCAAGCAGCAT 4365
Db 3694 CACAGTCGATGAATCCAGAAAGCGCCATTTTCCACCATCATATTCGGCAAGCAGCAT 3753
Qy 4366 GCCATAGGTCAAGCAGAGATCTTCGCGTCCGGCATCGCGCTTGAGCCTGGCGAACA 4425
Db 3754 GCCATAGGTCAAGCAGAGATCTTCGCGTCCGGCATCGCGCTTGAGCCTGGCGAACA 3813
Qy 4426 GTTCGCTGGCGGAGCCCTGATGCTCTTGTCCAGATCATCTCTGATCGACACAGCGG 4485
Db 3814 GTTCGCTGGCGGAGCCCTGATGCTCTTGTCCAGATCATCTCTGATCGACACAGCGG 3873
Qy 4486 TTTCATCCGAGTACGTCTGCTCGATGCGATGTTTCGCTGGTGGTGGTGGTGGTGG 4545
Db 3874 TTTCATCCGAGTACGTCTGCTCGATGCGATGTTTCGCTGGTGGTGGTGGTGGTGG 3933
Qy 4546 TAGCCGATCAAGCGTATGAGCGCCCGCATTCGATGAGCGATGATGATGATGATGATG 4605
Db 3934 TAGCCGATCAAGCGTATGAGCGCCCGCATTCGATGAGCGATGATGATGATGATGATG 3993
Qy 4606 CAGGAGCAAGTGAATGACAGGAGATCTTCGCGCGCAGCTTCGCCCAATAGCAGCAGT 4665
Db 3994 CAGGAGCAAGTGAATGACAGGAGATCTTCGCGCGCAGCTTCGCCCAATAGCAGCAGT 4053
Qy 4666 CCGTTCCCGCTTCACTGACAAAGTGAAGCAGCTGCGCAAGGAAACCGCGTGGGCCA 4725
Db 4054 CCGTTCCCGCTTCACTGACAAAGTGAAGCAGCTGCGCAAGGAAACCGCGTGGGCCA 4113
Qy 4726 GCCAGATAGCGCGCTCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4785
Db 4114 GCCAGATAGCGCGCTCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4173
Qy 4786 TGACAAAAGAACCGGGCGCCCTCGCTGACAGCGGAAACAGCGCGCATCAGAGCAGC 4845
Db 4174 TGACAAAAGAACCGGGCGCCCTCGCTGACAGCGGAAACAGCGCGCATCAGAGCAGC 4233
Qy 4846 CGATTGCTGTTGTCAGCTCATAGCCGATAGCTCTTCCACCAAGCGCGCGGGAAC 4905
Db 4234 CGATTGCTGTTGTCAGCTCATAGCCGATAGCTCTTCCACCAAGCGCGCGGGAAC 4293
Qy 4906 CTGCGTGAATCCATCTGTTCAATCATGCGAAACGATCCTCATCTGCTCTTGTATCAG 4965
Db 4294 CTGCGTGAATCCATCTGTTCAATCATGCGAAACGATCCTCATCTGCTCTTGTATCAG 4353
Qy 4966 ATCTTGATCCCTGCGCATCAGATCTTGGCGGCAAGAAAGCCATCAGTTACTTTGC 5025
Db 4354 ATCTTGATCCCTGCGCATCAGATCTTGGCGGCAAGAAAGCCATCAGTTACTTTGC 4413
```

```
Qy 5026 AGGCTTTCCAACTTACCAGAGGGCGCCCGCTGCGCAATTCGGTTTCGCTGCTGCTC 5085
Db 4414 AGGCTTTCCAACTTACCAGAGGGCGCCCGCTGCGCAATTCGGTTTCGCTGCTGCTC 4473
Qy 5086 ATAAACCCGCCAGTCTAGCTATCGCATGTAAGCCACTGCAAGCTACTGCTTCTCT 5145
Db 4474 ATAAACCCGCCAGTCTAGCTATCGCATGTAAGCCACTGCAAGCTACTGCTTCTCT 4533
Qy 5146 TTGCGCTTGGTTCCTTCCATAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCC 5205
Db 4534 TTGCGCTTGGTTCCTTCCATAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCC 4593
Qy 5206 ACCGTTTTCGCGACTGCTTCTAGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCG 5265
Db 4594 ACCGTTTTCGCGACTGCTTCTAGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCG 4653
Qy 5266 AGTGTCTTGGCGAGCGGTG 5283
Db 4654 AGTGTCTTGGCGAGCGGTG 4671

RESULT 2
AR260588
LOCUS 8349 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 16 from patent US 6489542.
ACCESSION AR260588
VERSION AR260588.1 GI:27311143
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 8349)
AUTHORS Corbin,D.R. and Romano,C.P.
TITLE Methods for transforming plants to express Cry2Ab
JOURNAL .delta.-endotoxins targeted to the plastids
Patent: US 6489542-A 16 03-DEC-2002;
FEATURES
Location/Qualifiers
source 1..8349
BASE COUNT 1959 a 2355 c 2088 g 1947 t
ORIGIN
Query Match 42.4%; Score 2240; DB 6; Length 8349;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 2409; Conservative 0; Mismatches 25; Indels 144; Gaps 1;

Qy 2706 CTAGACGTAATCATGTCATAGCTGTTTCCTGTTGAAATTCGTTATCCGCTCACAAATCC 2765
Db 5911 CTGCGGTAAATCATGTCATAGCTGTTTCCTGTTGAAATTCGTTATCCGCTCACAAATCC 5970
Qy 2766 ACACAACTACGAGCGGAGCATAAAGTGTAAAGCTGGGGTGCCTAATGAGTGAGCTA 2825
Db 5971 ACACAACTACGAGCGGAGCATAAAGTGTAAAGCTGGGGTGCCTAATGAGTGAGCTA 6030
Qy 2826 ACTCAATTAATGCGTTGCGTCTCATCTGCCCTTTCCAGTCCGGAAACCTGTCGTGCA 2885
Db 6031 ACTCAATTAATGCGTTGCGTCTCATCTGCCCTTTCCAGTCCGGAAACCTGTCGTGCA 6090
Qy 2886 GCTGCAATTAATCAATCGCCCAACGCGCGGAGAGCGGTTTCGTTATTCGGCGCTCTTC 2945
Db 6091 GCTGCAATTAATCAATCGCCCAACGCGCGGAGAGCGGTTTCGTTATTCGGCGCTCTTC 6150
Qy 2946 CGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCAGC 3005
Db 6151 CGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCAGC 6210
Qy 3006 TCACCTAAAGCGGTAATACGGTTATCCACAGAAATCAGGGATACGCGAGGAACAT 3065
Db 6211 TCACCTAAAGCGGTAATACGGTTATCCACAGAAATCAGGGATACGCGAGGAACAT 6270
Qy 3066 GTGAGCAAAAGCCAGCAAAAGCCAGGAACCGTAAAAAGCGCGGTTGCTCGCGCTTTT 3125
```


RESULT 3	SYN8KMRGC	2661 bp	DNA	linear	SYN 27-APR-1993
LOCUS	Synthetic plasmid pK18 (pBRNeo/pUC18)			kanamycin resistance gene,	
DEFINITION	complete cds.				
ACCESSION	M17626				
VERSION	M17626.1	GI:207845			
KEYWORDS	complete genome; kanamycin resistance.				
SOURCE	unidentified cloning vector				
ORGANISM	artificial sequences; vectors.				
REFERENCE	1 (bases 1 to 2661)				
AUTHORS	Pridmore,R.D.				
TITLE	New and versatile cloning vectors with kanamycin-resistance marker				
JOURNAL	Gene 56 (2-3), 309-312 (1987)				
MEDLINE	88056335				
PubMed	3315864				
COMMENT	Original				
FEATURES	source text: Plasmid pBRNeo and pUC18 DNA.				
source	Location/Qualifiers				
	1..2661				
	/organism="unidentified cloning vector"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:45136"				
	365..1159				
	/note="kanamycin"				
	/codon start=1				
	/transl table=11				
	/protein_id="AAA72612.1"				
	/db_xref="GI:207846"				
	/translation="MIEQDGLHAGSPAAWVERLFGVDWAQQTIGCSDAVFRLSAQR				
	PVLFKIDLSGALNEIQDEARLSWLAITGVCFAAVLDVVTEAGRWLLLGVEPQGL				
	LSHSLPAEKVSIMADMRHLHTDPAFCFPHQAKHRIERATERNAGLVQDDDLDE				
	EHQLGAPAEFLKAMPDGEDLVVTHGDACLPNIMVNGRFSGFIQGRGLGVADRY				
	QDIALATRDIAEELGGEWADRFLVLYGIAAPDSORIAFVRLLLDEFT"				
BASE COUNT	584 a 724 c 769 g 584 t				
ORIGIN					
	Query Match	42.38;	Score 2233.4;	DB 12;	Length 2661;
	Best Local Similarity	59.4%;	Pred. No. 0;		
	Matches 2403;	Conservative	0;	Mismatches	26;
				Indels	144;
				Gaps	1;
QY	2711	CGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATGTTATCCGGCTCACAAATCCACACA	2770		
Db	2443	CGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATGTTATCCGGCTCACAAATCCACACA	2384		
QY	2771	ACATACAGCGCGAAGCATAAAGTGTAAAGCCTGGGGTGCTTAATGAGTGAGTAACTCA	2830		
Db	2383	ACATACAGCGCGAAGCATAAAGTGTAAAGCCTGGGGTGCTTAATGAGTGAGTAACTCA	2324		
QY	2831	CATTAAATGGGTTCGGCTCACTGCCGCTTCCAGTTCGGGAAACCTGTGTGCGACCTGC	2890		
Db	2323	CATTAAATGGGTTCGGCTCACTGCCGCTTCCAGTTCGGGAAACCTGTGTGCGACCTGC	2264		
QY	2891	ATTAAATGAATCGGCAACAGCGCGGGAGAGCGGTTTGCGTATTGGCGCTCTTCGCGCTT	2950		
Db	2263	ATTAAATGAATCGGCAACAGCGCGGGAGAGCGGTTTGCGTATTGGCGCTCTTCGCGCTT	2204		
QY	2951	CCTCGCTCACTGACTCGCTCGCTCGGTTCGGCTTCGGCTGGCGGAGCGGTATCAGCTCACT	3010		
Db	2203	CCTCGCTCACTGACTCGCTCGCTCGGTTCGGCTTCGGCTGGCGGAGCGGTATCAGCTCACT	2144		
QY	3011	CAAAGCGGTAAATACGGTTATCCACAGAATCAGGGGATAACCGCAGGAAACATCTGAG	3070		
Db	2143	CAAAGCGGTAAATACGGTTATCCACAGAATCAGGGGATAACCGCAGGAAACATCTGAG	2084		
QY	3071	CAAAAGCCACGAAAAAGCCAGGAAACCGTAAAAAGCCGGTGTGCGGTTTTTCCATA	3130		
Db	2083	CAAAAGCCACGAAAAAGCCAGGAAACCGTAAAAAGCCGGTGTGCGGTTTTTCCATA	2024		
QY	3131	GGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTTCAGAGTGCAGAACCC	3190		
Db	2023	GGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTTCAGAGTGCAGAACCC	1964		

```
QY 4271 TATGTCCTGATAGCGGTCCGCCACACCCAGCGCGGCCACAGTCGATGAATCCAGAAAGCG 4330
Db 1027 TATGTCCTGATAGCGGTCCGCCACACCCAGCGCGGCCACAGTCGATGAATCCAGAAAGCG 368
QY 4331 GCCATTTTCCACCATGATATTCCGCAAGCAGGAGATCGCCATGGTTCACGACGAGATCCCTC 4390
Db 967 GCCATTTTCCACCATGATATTCCGCAAGCAGGAGATCGCCATGGTTCACGACGAGATCCCTC 308
QY 4391 GCCGTCGGGATGCGCGCTTGAGCTTGGCGAAACAGTTCCGCTGGCGGAGCCCTCGATG 4450
Db 907 GCCGTCGGGATGCGCGCTTGAGCTTGGCGAAACAGTTCCGCTGGCGGAGCCCTCGATG 848
QY 4451 CTCCTTCGTCAGATCATCTTCATCGCAAGACCGGCTTCATCGGAGTACGTCCTCGCTC 4510
Db 847 CTCCTTCGTCAGATCATCTTCATCGCAAGACCGGCTTCATCGGAGTACGTCCTCGCTC 788
QY 4511 GATCGGATGTTTCGCTTGGTGGTTCGATAGGCGAGTAGCGCGATCAAGCGTATGCGAGCG 4570
Db 787 GATCGGATGTTTCGCTTGGTGGTTCGATAGGCGAGTAGCGCGATCAAGCGTATGCGAGCG 728
QY 4571 CCGCATTTGCATCAGCCATGATGATATCTTTCTCGGAGGAGCAAGGTGATGATGACAGGAG 4630
Db 727 CCGCATTTGCATCAGCCATGATGATATCTTTCTCGGAGGAGCAAGGTGATGATGACAGGAG 668
QY 4631 ATCTGCGCCCGGCACTTCGCGCCCAATAGCAGCAGTCCCTTCCGCTTCAGTGACAACTGC 4690
Db 667 ATCTGCGCCCGGCACTTCGCGCCCAATAGCAGCAGTCCCTTCCGCTTCAGTGACAACTGC 608
QY 4691 GAGCAGAGTGGCGAAGAACCGCCGTCGTGGCGAGCAGCAGATAGCGGCGCTCGCTCGTC 4750
Db 607 GAGCAGAGTGGCGAAGAACCGCCGTCGTGGCGAGCAGCAGATAGCGGCGCTCGCTCGTC 548
QY 4751 CTGCAGTTTCATTCAGGCGACCGGACGTCGTCGTTTGACAAAAGAACCGGCGCCCTCG 4810
Db 547 TTGAGGTTTCATTCAGGCGACCGGACGTCGTCGTTTGACAAAAGAACCGGCGCCCTCG 488
QY 4811 CGTGAAGCGGAAACCGCGGCATCAGAGCAGCGGATGCTGTGTGTGTCGCGAGTCATA 4870
Db 487 CGTGAAGCGGAAACCGCGGCATCAGAGCAGCGGATGCTGTGTGTGTCGCGAGTCATA 428
QY 4871 GCCCAATAGCTCTCCACCCAGCGCGGAGAACCTGCGTGCAATCCATCTGTTCAAT 4930
Db 427 GCCCAATAGCTCTCCACCCAGCGCGGAGAACCTGCGTGCAATCCATCTGTTCAAT 368
QY 4931 CATCGGAAACGATCCTCATCTGCTCTTGATCAGATCTTGATCCCTCGGCGCATCAGAT 4990
Db 367 CATCGGAAACGATCCTCATCTGCTCTTGATCAGATCTTGATCCCTCGGCGCATCAGAT 308
QY 4991 CTTTGGCGGCAAGAAAGCCATCCAGTTTACTTTTGCAGGGCTTCCCAACCTTACGAGGG 5050
Db 307 CTTTGGCGGCAAGAAAGCCATCCAGTTTACTTTTGCAGGGCTTCCCAACCTTACGAGGG 248
QY 5051 CGCCCGCAGCTGGCAATTCGGGTCGCTTGCTGTCATATAAACCGCCAGTCTAGCTATCG 5110
Db 247 CGCCCGCAGCTGGCAATTCGGGTCGCTTGCTGTCATATAAACCGCCAGTCTAGCTATCG 188
QY 5111 CCATGTAAGCCCATGCAAGTACTCCTCTTCTTTGCGCTTTCGCTTTTCCCTTCTCCA 5170
Db 187 CCATGTAAGCCCATGCAAGTACTCCTCTTCTTTGCGCTTTCGCTTTTCCCTTCTCCA 128
QY 5171 GATAGCCAGTAGCTGACATTCATCCGGGCTCAGCACCGTTTCTGCGACTGGCTTCTTA 5230
Db 127 GATAGCCAGTAGCTGACATTCATCCGGGCTCAGCACCGTTTCTGCGACTGGCTTCTTA 68
QY 5231 CGTGTTCGCTTCTTTTTCAGAGCCCTTTCGCGCTTTCGCTTTCGCTTTCGCGAGCGTG 5283
Db 67 CGTGTTCGCTTCTTTTTCAGAGCCCTTTCGCGCTTTCGCTTTCGCGAGCGTG 15
```

RESULT 4

SYNCCDBA/c 2999 bp DNA linear SYN 16-MAR-2000
LOCUS
DEFINITION Cloning vector pKil118 ccdB gene, complete cds and

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
FEATURES
source

kanamycin-resistance (KmR) gene fragment.
L38498
L38498.1 GI:986977
ccdB gene; cloning vector; kanamycin resistance.
Cloning vector pKil118
Cloning vector pKil118
artificial sequences; vectors.
1 (bases 1 to 2999)
Bernard,P.
New ccdB positive-selection cloning vectors with kanamycin or
chloramphenicol selectable markers
Gene 162 (1), 159-160 (1995)
96009896
7557407

Location/Qualifiers
1..2999

/organism="Cloning vector pKil118"
/mol_type="genomic DNA"
/db_xref="taxon:118286"

1..381

/gene="ccdB"

1..381

/gene="ccdB"

/codon_start=1

/transl_table=11

/protein_id="AAC42103.1"

/db_xref="GI:986977"

/translation="MTWITNSSSVPGDPLSTCHASLSLQFKVYTKRSRVLKFDV
VQSDIIDTPRRMVIPLASARLLSDKVSRELYPVVHIGDESWRMVTTDMASVPVSVIG
BEVADLSHRENDIKNAINLFWGI"

BASE COUNT 670 a 799 c 860 g 670 t

Query Match 42.0%; Score 2218.4; DB 12; Length 2999;

Best Local Similarity 93.4%; Pred. No. 0;

Matches 2388; Conservative 0; Mismatches 26; Indels 144; Gaps 1;

QY 2726 AGCTGTTTCCTGTCGTGAAATTTTATCCGCTCAAAATTCACACAAATACGAGCCGAA 2785
Db 2999 AGCTGTTTCCTGTCGTGAAATTTTATCCGCTCAAAATTCACACAAATACGAGCCGAA 2940
QY 2786 GCATAAGTGTAAAGCTTGGGTGCTTAATGAGTGAGCTAACTCACAATTAATCGCTTC 2845
Db 2939 GCATAAGTGTAAAGCTTGGGTGCTTAATGAGTGAGCTAACTCACAATTAATCGCTTC 2880
QY 2846 GCTCACTCCCGCTTTCAGTCGGGAAACCTGTCGTCAGCTGCATTAAATGAATCGCC 2905
Db 2879 GCTCACTCCCGCTTTCAGTCGGGAAACCTGTCGTCAGCTGCATTAAATGAATCGCC 2820
QY 2906 AACGCGCGGAGAGGCGGTTTGGCTATTGGGCGCTCTTCGCTTCTCGCTCACTGACT 2965
Db 2819 AACGCGCGGAGAGGCGGTTTGGCTATTGGGCGCTCTTCGCTTCTCGCTCACTGACT 2760
QY 2966 CGCTCGCTCGCTGCTTCCGCTCGGCGAGCGGTATCAGCTCACTCAAGCGGTATAC 3025
Db 2759 CGCTCGCTCGCTGCTTCCGCTCGGCGAGCGGTATCAGCTCACTCAAGCGGTATAC 2700
QY 3026 GGTATTCCACAGAAATCAGGGGATTAACGAGGAAAGACATGTGAGCAAAAGCGCAGAAA 3085
Db 2699 GGTATTCCACAGAAATCAGGGGATTAACGAGGAAAGACATGTGAGCAAAAGCGCAGAAA 2640
QY 3086 AGGCGAGAAACCGTAAAGGCGCGGTGCTGGGCTTTTCCATAGAGCTCGCCCCCTG 3145
Db 2639 AGGCGAGAAACCGTAAAGGCGCGGTGCTGGGCTTTTCCATAGAGCTCGCCCCCTG 2580
QY 3146 ACGAGCATCAAAAAATCGAGCTCAAGTCAGAGTGGCGGAGAAACCCGACAGGACTATAA 3205
Db 2579 ACGAGCATCAAAAAATCGAGCTCAAGTCAGAGTGGCGGAGAAACCCGACAGGACTATAA 2520
QY 3206 GATACAGCGCTTCCCGCTTGGAAAGCTCCCTCGTGGCGCTCTCTGTTCCGACCTCGCGC 3265
Db 2519 GATACAGCGCTTCCCGCTTGGAAAGCTCCCTCGTGGCGCTCTCTGTTCCGACCTCGCGC 2460

ORGANISM Cloning vector pKIL119
REFERENCE 1 (bases 1 to 2999)
AUTHORS Bernard, P.
TITLE New ccdB positive-selection cloning vectors with kanamycin or chloramphenicol selectable markers
JOURNAL Gene 162 (1), 159-160 (1995)
MEDLINE 96009896
PUBMED 7557407
FEATURES
Location/Qualifiers
1..2999
/organism="Cloning vector pKIL119"
/mol_type="genomic DNA"
/db_xref="taxon:118287"
1..381
/gene="ccdB"
1..381
/gene="ccdB"
/codon_start=1
/transl_table=1
/protein_id="AAC42104.1"
/db_xref="GI:986980"
/translation="MTMITPSLHACRSTLEDPVPSSNSLFQKYYTKRSRVLFDV
VQSDIIDPGRWVPLASARILSKVSRRELYPVHIGDESWRMTTDMASVPVSVIG
BEVADLSHRENDIKNAIINFWGI"
BASE COUNT 672 a 801 c 857 g 669 t
ORIGIN
Query Match 42.0%; Score 2218.4; DB 12; Length 2999;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 2388; Conservative 0; Mismatches 26; Indels 144; Gaps 1;
QY 2726 AGCTGTTTCCTGTGGAATTTGTTATCCGCTCACAAATCCACACATACGACCGGAA 2785
DB |||||
QY 2786 GCATAAGTGAAGCCTGGGTGCTTAATGATGAGCTAACTACATTAATTCGTTGC 2845
DB |||||
QY 2939 GCATAAGTGAAGCCTGGGTGCTTAATGATGAGCTAACTACATTAATTCGTTGC 2880
DB |||||
QY 2846 GCTCACTGCCCTTCCAGTCGGGAACCTGTCGTGCGAGCTGATTAATGAATCGCC 2905
DB |||||
QY 2879 GCTCACTGCCCTTCCAGTCGGGAACCTGTCGTGCGAGCTGATTAATGAATCGCC 2820
DB |||||
QY 2906 AACCGCGGGGAGAGCGGTTTTCGCTATTGGCGCTCTTCCGCTTCCCTCGCTCACTGACT 2965
DB |||||
QY 2819 AACCGCGGGGAGAGCGGTTTTCGCTATTGGCGCTCTTCCGCTTCCCTCGCTCACTGACT 2760
DB |||||
QY 2966 CGCTGCGCTCGGTTCGCTGCGCGAGCGGTATCAGCTCACTCAAGGCGGTATATAC 3025
DB |||||
QY 2759 CGCTGCGCTCGGTTCGCTGCGCGAGCGGTATCAGCTCACTCAAGGCGGTATATAC 2700
DB |||||
QY 3026 GGTATCCACAGAAATCAGGCGGATACGAGGAAAGAAATGAGCAAAAGGCCAGCAAA 3085
DB |||||
QY 2699 GGTATCCACAGAAATCAGGCGGATACGAGGAAAGAAATGAGCAAAAGGCCAGCAAA 2640
DB |||||
QY 3086 AGCCAGGAACCGTAAAAAGGCGGCTGCTGCGGCTTTTCATAGGCTCGCCGCCCTG 3145
DB |||||
QY 2639 AGCCAGGAACCGTAAAAAGGCGGCTGCTGCGGCTTTTCATAGGCTCGCCGCCCTG 2580
DB |||||
QY 3146 ACAGCATCACAAAATCCAGCTCAAGTCAGAGTGGCGGAACCCGACAGGACTATAAA 3205
DB |||||
QY 2579 ACAGCATCACAAAATCCAGCTCAAGTCAGAGTGGCGGAACCCGACAGGACTATAAA 2520
DB |||||
QY 3206 GATACAGGCGTTTCCCTCGGAAGCTCCCTCGTGGCTCTCCTGTTCCGACCTCGCGC 3265
DB |||||
QY 2519 GATACAGGCGTTTCCCTCGGAAGCTCCCTCGTGGCTCTCCTGTTCCGACCTCGCGC 2460
DB |||||
QY 3266 TTACCGATACCTGTCGCGCTTTCCTCTCGGAGCGTGGCTTCTCTAGCTCAC 3325
DB |||||
QY 2459 TTACCGGATACCTGTCGCGCTTTCCTCTCGGAGCGTGGCTTCTCTAATGCTCAC 2400
DB |||||
QY 3326 GCTGTAGTATCTCAGTTCGGTGTAGTGTGTTCCGCTCCAGGCTGGGCTGTGTGACGAAC 3385
DB |||||

DB |||||
QY 2399 GCTGTAGTATCTCAGTTCGGTGTAGTTCGTCGCTCAAGCTGGGCTGTGTGCGCAAC 2340
QY |||||
QY 3386 CCCCCGTTCCAGCCCGACCGCTCGCTTATCCGGTAACTATCGTCTTGAGTCCAAACCCGG 3445
DB |||||
QY 2339 CCCCCGTTCCAGCCCGACCGCTCGCTTATCCGGTAACTATCGTCTTGAGTCCAAACCCGG 2280
QY |||||
QY 3446 TAAGACACGACTTATCGCCACTCGGACGAGCAGCTGCTGTAACAGGATTAGCAGAGCGAGT 3505
DB |||||
QY 2279 TAAGACACGACTTATCGCCACTCGGACGAGCAGCTGCTGTAACAGGATTAGCAGAGCGAGT 2220
QY |||||
QY 3506 ATGTAGCGGCTGTACAGAGTTCTTGAAGTGTGGCTTAACCTACGCTACACTAGAGAA 3565
DB |||||
QY 2219 ATGTAGCGGCTGTACAGAGTTCTTGAAGTGTGGCTTAACCTACGCTACACTAGAGAA 2160
QY |||||
QY 3566 CAGTATTTGGTATCTGCTGTGAAGCCAGTTACCTTCGGAAAGAGTTGGTAGCT 3625
DB |||||
QY 2159 CAGTATTTGGTATCTGCTGTGAAGCCAGTTACCTTCGGAAAGAGTTGGTAGCT 2100
QY |||||
QY 3626 CTTGATCCGGCAAAACAAACCGCTGTAGCGGTGTTTTTTTGTTCGAAGCAGCAGA 3685
DB |||||
QY 2099 CTTGATCCGGCAAAACAAACCGCTGTAGCGGTGTTTTTTTGTTCGAAGCAGCAGA 2040
QY |||||
QY 3686 TTACGGCGCAGAAAAAAGGATCTCAAGAAAGTCTTTTGTATCTTTCTACGGGCTGACG 3745
DB |||||
QY 2039 TTACGGCGCAGAAAAAAGGATCTCAAGAAAGTCTTTTGTATCTTTCTACGGGCTGACG 1980
QY |||||
QY 3746 CTCAGTGGAAACGAAACTCAGCTTAAGGATTTTGTCTATGAGATTATCGTCGACCAAG 3805
DB |||||
QY 1979 CTCAGTGGAAACGAAACTCAGCTTAAGGATTTTGTCTATGAGATTATCAAAAGGATCT 1920
QY |||||
QY 3806 CGGCCATCTGCTCCTCCACTCTCGAGTTCCGGGGCATGATGCGCGATAGCCGCTGC 3865
DB |||||
QY 1919 TCACCTAGATCCTT----- 1906
QY |||||
QY 3866 TGGTTTCCTGGATGCGACGAGATTTCACCTGCGGTAGAACTCCGCGAGGTGCTCCAGCC 3925
DB |||||
QY 1905 ----- 1906
QY |||||
QY 3926 TCAGGCGAGCTGAACCAACTCGCGAGGGATCGAGCCCGGGTGGCGAAGAACTCCA 3985
DB |||||
QY 1905 -----TTGGGTGGCGAAGAACTCCA 1884
QY |||||
QY 3986 GCATGAGATCCCGCGCTGAGGATCATCCAGCCGCGTCCCGGAAAAAGATTCGGAAGC 4045
DB |||||
QY 1883 GCATGAGATCCCGCGCTGAGGATCATCCAGCCGCGTCCCGGAAAAAGATTCGGAAGC 1824
QY |||||
QY 4046 CCAACCTTTTCATAGAGGCGGCTGGAATCGAAATCTCGTATGCGAGCTTGGGCGTGC 4105
DB |||||
QY 1823 CCAACCTTTTCATAGAGGCGGCTGGAATCGAAATCTCGTATGCGAGCTTGGGCGTGC 1764
QY |||||
QY 4106 CTTTGGTGGTCAATTCGAAACCCCGAGAGTCCCGCTCAGAAAGAACTCGTCAAGAGGCGATA 4165
DB |||||
QY 1763 CTTTGGTGGTCAATTCGAAACCCCGAGAGTCCCGCTCAGAAAGAACTCGTCAAGAGGCGATA 1704
QY |||||
QY 4166 GAGGGGATGCGCTCGGAATCGGGAGCGGATACCGTAAAGCAGAGGAAAGCGGTACG 4225
DB |||||
QY 1703 GAGGGGATGCGCTCGGAATCGGGAGCGGATACCGTAAAGCAGAGGAAAGCGGTACG 1644
QY |||||
QY 4226 CCATTTCGCGCCCAAGCTCTTCAGCAATATCAGGGTAGCAAGCTATGCTCCTGATAGCG 4285
DB |||||
QY 1643 CCATTTCGCGCCCAAGCTCTTCAGCAATATCAGGGTAGCAAGCTATGCTCCTGATAGCG 1584
QY |||||
QY 4286 GTCGCGCACACCCAGCGGCGCAGTCGATGAATCCAGAAAGCGGCAATTTCCACCAT 4345
DB |||||
QY 1583 GTCGCGCACACCCAGCGGCGCAGTCGATGAATCCAGAAAGCGGCAATTTCCACCAT 1524
QY |||||
QY 4346 GATATTTCGCAAGCAGGCAATCGCATGCGGTACGAGAGATCTCGCCGTCGCGGCAATGCG 4405
DB |||||
QY 1523 GATATTTCGCAAGCAGGCAATCGCATGCGGTACGAGAGATCTCGCCGTCGCGGCAATGCG 1464
QY |||||
QY 4406 CGCCTTAGGCTCGCGCAAGTTCGGCTGGCGGAGCCCTGATGCTCTTCGTCGACATC 4465
DB |||||

Db 1463 CGCCTTGAGCTGCGCAACAGTTGGCTGGCGGAGCCCTGATGCTCTTCGTCCAGATC 1404
Qy 4466 ATCTGTATCGAAGACCGGCTTCCATCCGAGTACGTGCTCGTTCGATGCGATGTTTCGC 4525
Db 1403 ATCTGTATCGAAGACCGGCTTCCATCCGAGTACGTGCTCGTTCGATGCGATGTTTCGC 1344
Qy 4526 TTGTTGTGCGAATGGGAGGTAGCGGATCAAGGATGAGCGGCGCGGATTCATCAGC 4585
Db 1343 TTGTTGTGCGAATGGGAGGTAGCGGATCAAGGATGAGCGGCGCGGATTCATCAGC 1284
Qy 4586 CATGATCGATCTTCTCGGAGGAGCAAGTGAGATGACAGAGATCCTGCCCGCGCAC 4645
Db 1283 CATGATCGATCTTCTCGGAGGAGCAAGTGAGATGACAGAGATCCTGCCCGCGCAC 1224
Qy 4646 TTGCGCCAAATAGCAGCGGATCCCTTCCGCTTTCAGTGACAACTCGAGCAGCTCGCA 4705
Db 1223 TTGCGCCAAATAGCAGCGGATCCCTTCCGCTTTCAGTGACAACTCGAGCAGCTCGCA 1164
Qy 4706 AGGAACCGCGTGTGCGGAGGAGCAAGTGAGATGACAGAGATCCTGCCCGCGCAC 4765
Db 1163 AGGAACCGCGTGTGCGGAGGAGCAAGTGAGATGACAGAGATCCTGCCCGCGCAC 1104
Qy 4766 GGCAACCGGAGTGTGCTTTCGCAAAAGAAACCGGCGCGCTCGGCTGACAGCGGAA 4825
Db 1103 GGCAACCGGAGTGTGCTTTCGCAAAAGAAACCGGCGCGCTCGGCTGACAGCGGAA 1044
Qy 4826 CAGCGCGGATCAGAGAGCGGATGTCTGTTGCGGAGTATGAGCGGAGTACGCTCTC 4885
Db 1043 CAGCGCGGATCAGAGAGCGGATGTCTGTTGCGGAGTATGAGCGGAGTACGCTCTC 984
Qy 4886 CACCAAGCGCGGAGAACTGGTGCAATTCATCTTGTTCATCATGCGAAACGATCC 4945
Db 983 CACCAAGCGCGGAGAACTGGTGCAATTCATCTTGTTCATCATGCGAAACGATCC 924
Qy 4946 TCATCTGTCTTGTATCAGATCTTGTATCCCTGCGGCAATCATGCTTGGCGGCAAGAA 5005
Db 923 TCATCTGTCTTGTATCAGATCTTGTATCCCTGCGGCAATCATGCTTGGCGGCAAGAA 864
Qy 5006 AGCCATCCAGTTACTTTGAGAGGCTTCCCACTTACAGAGGCGGCGGCGGAGTACG 5065
Db 863 AGCCATCCAGTTACTTTGAGAGGCTTCCCACTTACAGAGGCGGCGGCGGAGTACG 804
Qy 5066 TTGCGGTTGCTGCTGCTCAAAACCGGCGGAGTCTAGCTATGCGCATGTAAAGCCACT 5125
Db 803 TTGCGGTTGCTGCTGCTCAAAACCGGCGGAGTCTAGCTATGCGCATGTAAAGCCACT 744
Qy 5126 GCAAGCTACCTGCTTCTTTCGCTTGGCTTGGCTTCCCTGTCAGATGCGGAGTACCT 5185
Db 743 GCAAGCTACCTGCTTCTTTCGCTTGGCTTGGCTTCCCTGTCAGATGCGGAGTACCT 684
Qy 5186 GACATTCATCGGCGTCAAGCAGGCTTCTGCGGAGTGGCTTTCAGCTGTTCCGCTTCT 5245
Db 683 GACATTCATCGGCGTCAAGCAGGCTTCTGCGGAGTGGCTTTCAGCTGTTCCGCTTCT 624
Qy 5246 TTAGAGGCTTGGCGGCTGAGTCTTGGCGGAGGCTG 5283
Db 623 TTAGAGGCTTGGCGGCTGAGTCTTGGCGGAGGCTG 586

RESULT 6
LOCUS AY189826/c 2604 bp DNA circular SYN 09-APR-2003
DEFINITION His-3 integration vector pJHAM001, complete sequence.
ACCESSION AY189826
VERSION AY189826.1 GI:28435536
KEYWORDS
SOURCE his-3 integration vector pJHAM001
ORGANISM his-3 integration vector pJHAM001
REFERENCE 1 (bases 1 to 2604)
AUTHORS Lee, D.W., Haag, J.R. and Aramayo, R.
TITLE Construction of strains for rapid homokaryon purification after integration of constructs at the histidine-3 (his-3) locus of

Neurospora crassa
Curr. Genet. 43 (1), 17-23 (2003)
12684841
2 (bases 1 to 2604)
Lee, D.W., Haag, J.R. and Aramayo, R.
Direct Submission
Submitted (03-DEC-2002) Biology, Texas A&M University, BSBW 415,
College Station, TX 77843-3258, USA
Location/Qualifiers
1. 2604
/organism="his-3 integration vector pJHAM001"
/mol_type="genomic DNA"
/db_xref="taxon:220724"
/note="modified pk18 vector"
2542..2572
/note="modified multi-cloning site"
BASE COUNT 561 a 714 c 760 g 569 t
ORIGIN
Query Match 41.2%; Score 2177.4; DB 12; Length 2604;
Best Local Similarity 93.2%; Pred. No. 0;
Matches 2359; Conservative 0; Mismatches 26; Indels 146; Gaps 2;
Qy 2753 CGCTCACAAATCCACAAACATACGAGCGGAGCATAAAGTGTAAAGCCTGGGTGCT 2812
Db 2542 CGCTCACAAATCCACAAACAT--AGCCGAGCATAAAGTGTAAAGCCTGGGTGCT 2485
Qy 2813 AATGAGTGAATTAATCACTTAATTAATGCTTGGCTGCTCACTGCCGCTTTCAGTCGGAA 2872
Db 2484 AATGAGTGAATTAATCACTTAATTAATGCTTGGCTGCTCACTGCCGCTTTCAGTCGGAA 2425
Qy 2873 ACCTGTCGTGCGAGTCATTAATGATCGGCAACGCGGCGGAGAGCGGCTTTCGCTA 2932
Db 2424 ACCTGTCGTGCGAGTCATTAATGATCGGCAACGCGGCGGAGAGCGGCTTTCGCTA 2365
Qy 2933 TTGCGCGCTCTTCGCGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCGC 2992
Db 2364 TTGCGCGCTCTTCGCGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCGC 2305
Qy 2993 GAGCGGTATCAGTCACTCAAGGGGTAATACGGTTATCCACAGATCAGGGATACG 3052
Db 2304 GAGCGGTATCAGTCACTCAAGGGGTAATACGGTTATCCACAGATCAGGGATACG 2245
Qy 3053 CAGGAAAGAAATGTGAGCAAAAGCCAGCAAAAGGCGGAGAAACCGTAAAGAGCGCGCT 3112
Db 2244 CAGGAAAGAAATGTGAGCAAAAGCCAGCAAAAGGCGGAGAAACCGTAAAGAGCGCGCT 2185
Qy 3113 TGCTGGCGCTTTTCCATAGGCTCCGCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 3172
Db 2184 TGCTGGCGCTTTTCCATAGGCTCCGCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 2125
Qy 3173 GTCAGAGTGGCGAAACCCGAGCAGACTATAAGATACAGGCGCTTCCCGCTCGAAGCT 3232
Db 2124 GTCAGAGTGGCGAAACCCGAGCAGACTATAAGATACAGGCGCTTCCCGCTCGAAGCT 2055
Qy 3233 CCCTGCTCGCTCTCTCTTCCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCTCC 3292
Db 2054 CCCTGCTCGCTCTCTCTTCCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCTCC 2005
Qy 3293 CTTGCGGAGGCTGGCGCTTCTCATAGCTCACGCTGAGTATCTCAGTTCGCTGTAGG 3352
Db 2004 CTTGCGGAGGCTGGCGCTTCTCATAGCTCACGCTGAGTATCTCAGTTCGCTGTAGG 1945
Qy 3353 TCGTTTCGCTCCAAAGTGGGCTGTGTGCAAGAACCCCGCTTTCAGCGCGGCTGCGCT 3412
Db 1944 TCGTTTCGCTCCAAAGTGGGCTGTGTGCAAGAACCCCGCTTTCAGCGCGGCTGCGCT 1885
Qy 3413 TATCGGTAATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3472
Db 1884 TATCGGTAATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1825
Qy 3473 CAGCCACTGGTAAACAGGATTAGCAGAGCGGATGTAGGCGGTGCTACAGAGTTCTTGA 3532

1824 CAGCCACTGGTAAACAGGATTAGCAGCGAGGATGTAGCGGTCTACAGAGTTCTTGA 1765
Db
3533 AGTGTGGCTTAACCTACGGCTACACTAGAGAAAGATTTGGTATCTGGCTCTGCTGA 3592
Qy
1764 AGTGTGGCTTAACCTACGGCTACACTAGAGAAAGATTTGGTATCTGGCTCTGCTGA 1705
Db
3593 AGCCAGTTACCTTCGGAAAGAGTTGGTATCTGGTATCTGGTATCTGGTATCTGGT 3652
Qy
1704 AGCCAGTTACCTTCGGAAAGAGTTGGTATCTGGTATCTGGTATCTGGTATCTGGT 1645
Db
3653 GTAGCGGTGGTATCTGGTATCTGGTATCTGGTATCTGGTATCTGGTATCTGGTATCT 3712
Qy
1644 GTAGCGGTGGTATCTGGTATCTGGTATCTGGTATCTGGTATCTGGTATCTGGTATCT 1585
Db
3713 AGATCCTTTGATCTTTCTACGGGTCTGAGCTCAGTGGAAAGAAAGTAAAG 3772
Qy
1584 AGATCCTTTGATCTTTCTACGGGTCTGAGCTCAGTGGAAAGAAAGTAAAG 1525
Db
3773 GGATTTGTCATGAGTATTCGTGCAACAAAGCGGCGCATCTGCTCCCTCCCTGCA 3832
Qy
1524 GGATTTGTCATGAGTATTCGTGCAACAAAGCGGCGCATCTGCTCCCTCCCTGCA 1478
Db
3833 GTTCGGGGCATGGATGCGCGGATAGCCGCTGCTGTTCTTCTGGATGCGAGGATTCG 3892
Qy
1477 ----- 1478
Db
3893 ACTCGCGGTAGAACTCCGCGAGGTGTCAGGCTCAGGCGAGTGAACCAACTCGCGA 3952
Qy
1477 ----- 1478
Db
3953 GGGGATCGAGCCCGGGTGGGAGAACTCCAGCATGAGATCCCGCGCTGAGGATCA 4012
Qy
1477 ----- TTGGGGTGGGCGAGAACTCCAGCATGAGATCCCGCGCTGAGGATCA 1429
Db
4013 TCAGCGCGGTCCCGGAAACCAATCCGAAGCCCAACCTTTCTAGAAAGCGCGGTGG 4072
Qy
1428 TCAGCGCGGTCCCGGAAACCAATCCGAAGCCCAACCTTTCTAGAAAGCGCGGTGG 1369
Db
4073 ANTGAAATCTGTGATGGCAGGTTGGGCTGCTTGGTGGCTCATTTGACACCCAGAG 4132
Qy
1368 AATCGAAATCTGTGATGGCAGGTTGGGCTGCTTGGTGGCTCATTTGACACCCAGAG 1309
Db
4133 TCCCGCTCAGAAAGTCTGTCAAGAAAGGAGATAGAAAGGAGTGGCTGCGAATCGGGAGC 4192
Qy
1308 TCCCGCTCAGAAAGTCTGTCAAGAAAGGAGATAGAAAGGAGTGGCTGCGAATCGGGAGC 1249
Db
4193 GGGGATACCGTAAAGCAAGAGGAGGTTGGGCTGCTTGGTGGCTCATTTGACACCCAGAG 4252
Qy
1248 GGGGATACCGTAAAGCAAGAGGAGGTTGGGCTGCTTGGTGGCTCATTTGACACCCAGAT 1189
Db
4253 ATCAGCGGTAGCCAAAGCTATGCTGATAGCGGTGCGGACACACCCAGCGGCGCAGTC 4312
Qy
1188 ATCAGCGGTAGCCAAAGCTATGCTGATAGCGGTGCGGACACACCCAGCGGCGCAGTC 1129
Db
4313 GATGAATCCAGAAAGCGGCATTTTCCACATGATATTCGGCAAGCAGGATTCGCCATG 4372
Qy
1128 GATGAATCCAGAAAGCGGCATTTTCCACATGATATTCGGCAAGCAGGATTCGCCATG 1069
Db
4373 GGTCAAGCAGGATCTCGCGTGGGATCGGCGCTTGGCTGCGGCGGCGGCGGCGGCGG 4432
Qy
1068 GGTCAAGCAGGATCTCGCGTGGGATCGGCGCTTGGCTGCGGCGGCGGCGGCGGCGG 1009
Db
4433 TGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4492
Qy
1008 TGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 949
Db
4493 CCGAGTAGCTGCTCGCTGATGCGATGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4552
Qy
948 CCGAGTAGCTGCTCGCTGATGCGATGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 889
Db
4553 ATCAAGCGTATGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 4612
Qy
888 ATCAAGCGTATGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 829
Db

Qy 4613 AAGGTGAGATGACAGGAGATCTGCGCCCGCACTTGGCCCAATAGCAGCAGTCCCTTCC 4672
Db 828 AAGGTGAGATGACAGGAGATCTGCGCCCGCACTTGGCCCAATAGCAGCAGTCCCTTCC 769
Qy 4673 CGTTTCAGTGAACAAGTCGAGCAGCTGCGCAAGAGAGCGCGTGGTGGCGAGCCAGA 4732
Db 768 CGTTTCAGTGAACAAGTCGAGCAGCTGCGCAAGAGAGCGCGTGGTGGCGAGCCAGA 709
Qy 4733 TAGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4792
Db 708 TAGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 649
Qy 4793 AAGAAACGGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4852
Db 648 AAGAAACGGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 589
Qy 4853 CTGTTGTGCCAGTATAGCCGAATAGCTCTCCACCCCAAGCGGCGGAGAACCTCGGTG 4912
Db 588 CTGTTGTGCCAGTATAGCCGAATAGCTCTCCACCCCAAGCGGCGGAGAACCTCGGTG 529
Qy 4913 CAATCCATCTTTGTTCAATCATGCGAAACGATCTCTCTCTCTGATCAGATCTTGA 4972
Db 528 CAATCCATCTTTGTTCAATCATGCGAAACGATCTCTCTCTCTGATCAGATCTTGA 469
Qy 4973 TCCCTCGCGCATCAGATCTTGGCGGCAAGAAAGCAATCCAGTTTACCTTTCAGGCGCT 5032
Db 468 TCCCTCGCGCATCAGATCTTGGCGGCAAGAAAGCAATCCAGTTTACCTTTCAGGCGCT 409
Qy 5033 CCCAACCTTACCAGAGGCGGCCCAAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5092
Db 408 CCCAACCTTACCAGAGGCGGCCCAAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 349
Qy 5093 CGCCAGTCTAGCTATCGCCATGTAAGCCCACTGCAAGCTACCTGCTTCTCTTTCGCT 5152
Db 348 CGCCAGTCTAGCTATCGCCATGTAAGCCCACTGCAAGCTACCTGCTTCTCTTTCGCT 289
Qy 5153 TGCCTTTTCCCTTGTCCAGATAGCCAGTAGCTGACATTCATCCGGGTCAGACCGTTT 5212
Db 288 TGCCTTTTCCCTTGTCCAGATAGCCAGTAGCTGACATTCATCCGGGTCAGACCGTTT 229
Qy 5213 CTGGGATCGCTTCTACGTTTCCGCTTCCCTTTCAGGCGCTTGGCGCTTTCAGTCT 5272
Db 228 CTGGGATCGCTTCTACGTTTCCGCTTCCCTTTCAGGCGCTTGGCGCTTTCAGTCT 169
Qy 5273 GCGGCGAGCTG 5283
Db 168 GCGGCGAGCTG 158

RESULT 7
AY189827/c

LOCUS 11373 bp DNA circular SYN 09-APR-2003
DEFINITION His-3 integration vector pJHAM002, complete sequence.

ACCESSION AY189827

VERSION AY189827.1 GI:28435537

KEYWORDS

SOURCE his-3 integration vector pJHAM002

ORGANISM his-3 integration vector pJHAM002

REFERENCE 1 (bases 1 to 11373)

AUTHORS Lee, D.W., Haag, J.R., and Aramayo, R.

TITLE Construction of strains for rapid homokaryon purification after

integration of constructs at the histidine-3 (his-3) locus of

Neurospora crassa

JOURNAL Curr. Genet. 43 (1), 17-23 (2003)

PUBMED 12684841

REFERENCE 2 (bases 1 to 11373)

AUTHORS Lee, D.W., Haag, J.R., and Aramayo, R.

TITLE Direct Submission

JOURNAL Submitted (03-DEC-2002) Biology, Texas A&M University, BSBW 415,

College Station, TX 77843-3258, USA

Location/Qualifiers

FEATURES


```
/gene="aph(3')-IIa"
/notes="aminoglycoside phosphotransferase"
/codon_start=1
/transl_table=11
/product="aph(3')-IIa"
/protein_id="AAO66595.1"
/db_xref="GI:29242890"
/translation="MIEQDGLHAGSPAVERLFGYDMAQQTIGCSDAAVFLSAGGR
LVSEKTDLSGALNELQDEARLSWLAITGVCPCAAVLDDVTEAGRWLLIGVPGDGL
PSSHAPAEKVISIMADARRLITDPATCFDQAKHRIERARTRMEAGLVDDDDIDE
HQGLAPAEILPARLKAMPDGEDLVTHGDACLPMWVNGRPSGIDCGRLGVADRY
QDIALNRDIAEELGGEWADRFVLVGIAPDPSQRIAFYRLLEDF"
misc_feature
1760..1858
/notes="RP4-mob region"
rep_origin
2621..2623
/notes="oriV"
BASE COUNT 665 a 844 c 872 g 655 t
ORIGIN
Query Match 39.5%; Score 2088.4; DB 12; Length 3036;
Best Local Similarity 86.7%; Pred. No. 0;
Matches 2434; Conservative 0; Mismatches 141; Indels 231; Gaps 3;
QY 2709 CACGTAATCATGGTCATAGCTGTTTCCTGTGTGAATTTGTTATCCGCTCAAAATCCACA 2768
DB |||
DB 3036 GGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAATTTGTTATCCGCTCAAAATCCACA 2977
QY 2769 CAACATACGACGCGGAAGCATTAAGTGTAAAGCCTGGGTGCTTAAGTGTAGTGAAGT 2828
DB |||
DB 2976 CAACATACGACGCGGAAGCATTAAGTGTAAAGCCTGGGTGCTTAAGTGTAGTGAAGT 2917
QY 2829 CACATTAATTCGTTGGCTCACTGCCGCTTTCCAGTCGGGAACCTGTCTGTCAGCT 2888
DB |||
DB 2916 CACATTAATTCGTTGGCTCACTGCCGCTTTCCAGTCGGGAACCTGTCTGTCAGCT 2857
QY 2889 GCATTAATGAATTCGCCCAACGCGGCGGAGAGCGGTTTGGTATTCGGGCTCTTCGCG 2948
DB |||
DB 2856 GCATTAATGAATTCGCCCAACGCGGCGGAGAGCGGTTTGGTATTCGGGCTCTTCGCG 2797
QY 2949 TTCCTCGCTCACTGACTCGCTCGCTCGCTGCTGCTGCGGCGGAGCGGTATCAGCTCA 3008
DB |||
DB 2796 TTCCTCGCTCACTGACTCGCTCGCTCGCTGCTGCTGCGGCGGAGCGGTATCAGCTCA 2737
QY 3009 CTCAAAGCGGTAATACCGGTTATCCACAGAAATCAGGGGATTAACGAGGAAAGAACATGTG 3068
DB |||
DB 2736 CTCAAAGCGGTAATACCGGTTATCCACAGAAATCAGGGGATTAACGAGGAAAGAACATGTG 2677
QY 3069 AGCAAAAGGCGAGCAAAAGGCGAGAACCGTAAAGAGCCGCTTGTGGCGTTTTCCTCA 3128
DB |||
DB 2676 AGCAAAAGGCGAGCAAAAGGCGAGAACCGTAAAGAGCCGCTTGTGGCGTTTTCCTCA 2617
QY 3129 TAGGCTTCGCGCCCTCGACGACATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 3188
DB |||
DB 2616 TAGGCTTCGCGCCCTCGACGACATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 2557
QY 3189 CCGCAGAGACTATAAGATACAGCGGTTTCCTGGAAGCTCCCTCGTGGCGCTCTCC 3248
DB |||
DB 2556 CCGCAGAGACTATAAGATACAGCGGTTTCCTGGAAGCTCCCTCGTGGCGCTCTCC 2497
QY 3249 TGTTCGACCCCTCGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGC 3308
DB |||
DB 2496 TGTTCGACCCCTCGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGC 2437
QY 3309 GCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGTTAGGTGCTTCCTCAAGCT 3368
DB |||
DB 2436 GCTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCCGTTAGGTGCTTCCTCAAGCT 2377
QY 3369 GGCTGTGTGCGAGAACCCCGCTTCAGCCGAGCGCTGGCGCTTATCCGTTACTATCG 3428
DB |||
DB 2376 GGCTGTGTGCGAGAACCCCGCTTCAGCCGAGCGCTGGCGCTTATCCGTTACTATCG 2317
QY 3429 TCTTGTAGTCCAAACCCGGTAAGACAGACTTATTCGCACTTGGCAGCAGCACTTGGTAACAG 3488
DB |||
```

QY 4338 TCACCATGATATTCGGCAGCAGGATCGCCATCGGTGTCACGACGATCTCTCGCGTCG 4397
Db 1176 TCACCATGATATTCGGCAGCAGGATCGCCATCGGTGTCACGACGATCTCTCGCGTCG 1117
QY 4398 GGCATCGCGCCTTGGCTGCGGAAAGTTCGGCTGCGGAGCCCTGATGCTCTTCG 4457
Db 1116 GGCATCGCGCCTTGGCTGCGGAAAGTTCGGCTGCGGAGCCCTGATGCTCTTCG 1057
QY 4458 TCACATCATCTGATCGACAGACCGGGTTCATCCGAGTACGTCGTCGATCGCA 4517
Db 1056 TCACATCATCTGATCGACAGACCGGGTTCATCCGAGTACGTCGTCGATCGCA 997
QY 4518 TGTTCGCTGCTGCTGATCGGAGTTCAGGATCAAGGATGACGAGCCGCGATT 4577
Db 996 TGTTCGCTGCTGCTGATCGGAGTTCAGGATCAAGGATGACGAGCCGCGATT 937
QY 4578 GCATCAGCATGATGATATCTTCGCGAGGACAGGTGAGATGACGAGATCTCTGC 4637
Db 936 GCATCAGCATGATGATATCTTCGCGAGGACAGGTGAGATGACGAGATCTCTGC 877
QY 4638 CCGGCACCTTCGCCAATAGCAGCCAGTCCCTTCGCGTTCAGTGACAAAGTCGACACA 4697
Db 876 CCGGCACCTTCGCCAATAGCAGCCAGTCCCTTCGCGTTCAGTGACAAAGTCGACACA 817
QY 4698 GCTCGCAGGAACGCCGCTGCTGCGCAGCAGATAGCCGCTGCTGCTGCTGCAAT 4757
Db 816 GCTCGCAGGAACGCCGCTGCTGCGCAGCAGATAGCCGCTGCTGCTGCTGCAAT 757
QY 4758 TCATTCAGGACACCGGACAGTTCGTTGACAAAAGAACCGGGCGCCCTGCGCTGAC 4817
Db 756 TCATTCAGGACACCGGACAGTTCGTTGACAAAAGAACCGGGCGCCCTGCGCTGAC 697
QY 4818 AGCGGAACACCGGGGATCAGACGAGCCGATTCGTTGTTGTCGCCAGTCATGCGCAAT 4877
Db 696 AGCGGAACACCGGGGATCAGACGAGCCGATTCGTTGTTGTCGCCAGTCATGCGCAAT 637
QY 4878 AGCTCTCCACCCAGCGGCGGAGAACCTGCGTGCATCAATCTGTTCAATCATGCGA 4937
Db 636 AGCTCTCCACCCAGCGGCGGAGAACCTGCGTGCATCAATCTGTTCAATCATGCGA 577
QY 4938 AACGATCTCATCTGCTCTTGTATCAGATTCGATCCCTGCGCATCAGATCTTGGC 4997
Db 576 AACGATCTCATCTGCTCTTGTATCAGATTCGATCCCTGCGCATCAGATCTTGGC 517
QY 4998 GGCAGGAAGCCATCCAGTTTACTTTTCAGGGCTTCCCAACCTTACAGAGGCGGCCCA 5057
Db 516 GGCAGGAAGCCATCCAGTTTACTTTTCAGGGCTTCCCAACCTTACAGAGGCGGCCCA 457
QY 5058 GCTGGCAATTCGGTTGCTGTCATGATGATGATGATGATGATGATGATGATGATGAT 5117
Db 456 GCTGGCAATTCGGTTGCTGTCATGATGATGATGATGATGATGATGATGATGATGAT 397
QY 5118 AGCCCACTGCAAGTACTGCTTCTCTTTGCGCTTTCGCTTTTCCCTTTCAGATAGCC 5177
Db 396 AGCCCACTGCAAGTACTGCTTCTCTTTGCGCTTTCGCTTTTCCCTTTCAGATAGCC 337
QY 5178 CAGTAGTGATTCATTCGCGGTCAGACCGTTTTCGCGATGCGTTTTCAGTGTTTC 5237
Db 336 CAGTAGTGATTCATTCATTCGCGGTCAGACCGTTTTCGCGATGCGTTTTCAGTGTTTC 277
QY 5238 CGCTTCTTTAGCAGCCCTTTCGCGCTGAGTTCGCTGCGGAGCGTG 5283
Db 276 CGCTTCTTTAGCAGCCCTTTCGCGCTGAGTTCGCTGCGGAGCGTG 231

RESULT 10
AY219859/c
LOCUS
DEFINITION Cloning vector pK19MECA2, complete sequence.
ACCESSION AY219859
VERSION AY219859.1 GI:29164979
KEYWORDS

SOURCE
ORGANISM Cloning vector pK19MECA2
REFERENCE Cloning vector pK19MECA2
1 (bases 1 to 3210)
AUTHORS Kirchner, O. and Tauch, A.
TITLE Tools for genetic engineering in the amino acid-producing bacterium
Corynebacterium glutamicum
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3210)
AUTHORS Kirchner, O. and Tauch, A.
TITLE Direct Submission
JOURNAL Submitted (15-JAN-2003) Department of Genetics, University of
Bielefeld, Universitaetstrasse 25, Bielefeld D-33615, Germany
FEATURES
Location/Qualifiers
1..3210
/organism="Cloning vector pK19MECA2"
/mol_type="genomic DNA"
/db_xref="taxon:224628"
755..1549
/gene="aph(3')-IIa"
755..1549
/gene="aph(3')-IIa"
/note="aminoglycoside phosphotransferase"
/codon_start=1
/transl_table=11
/product="aph(3')-IIa"
/protein_id="AA065226.1"
/db_xref="GI:29164980"
translation="MIEQGLHAGSPAAWVERLFQYDWAQQTIGSCDAVFLRLSAQGR
PVLFLKLSALNELQDEARLTLWATTCVCAAVLDVTTAGRDMLLLGVEPQDL
LSSHLAPAEKVISIMADAMRLHTLDPATCPFDHQAHRIRERATMEAGLVDDDLDE
BHQGLAPAEELPARLKAEMPDGDLVVTGHDACLNIWVENGFRSGFIDGRLGVADRY
ODIALATRDIAELGGEWADRLVLYIAAPDSQRIAFYRLDLEFF"
1934..2032
misc_feature
/note="RP4-mob region"
2795..2797
rep_origin
/note="oriV"
BASE COUNT 705 a 887 c 919 g 699 t
ORIGIN
Query Match 39.5%; Score 2088.4; DB 12; Length 3210;
Best Local Similarity 86.7%; Pred. No. 0;
Matches 2434; Conservative 0; Mismatches 141; Indels 231; Gaps 3;
QY 2709 GACGTAATCATGTCATAGCTGTTTCCTGTGTAATTTGTTATCGCTCACAATTCACA 2768
Db 3210 GCGTAATCATGTCATAGCTGTTTCCTGTGTAATTTGTTATCGCTCACAATTCACA 3151
QY 2769 CAACATACGAGCGGAGCAATAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGTAACT 2828
Db 3150 CAACATACGAGCGGAGCAATAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGTAACT 3091
QY 2829 CACATTAATTCGGTTCGCTCACTCCCGCTTTCAGTCGGGAAACCTGTGTCAGCT 2888
Db 3090 CACATTAATTCGGTTCGCTCACTCCCGCTTTCAGTCGGGAAACCTGTGTCAGCT 3031
QY 2889 GCATTAATGAATCGCCCAACCGCGGAGAGCGGTTTGGCTATTCGGCGCTCTTCGCG 2948
Db 3030 GCATTAATGAATCGCCCAACCGCGGAGAGCGGTTTGGCTATTCGGCGCTCTTCGCG 2971
QY 2949 TTCCTCGCTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3008
Db 2970 TTCCTCGCTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2911
QY 3009 CTCAGAGCGGTAATACGTTTATCCAGATCAGGGGATAACCCAGGAAAGACATGTG 3068
Db 2910 CTCAGAGCGGTAATACGTTTATCCAGATCAGGGGATAACCCAGGAAAGACATGTG 2851
QY 3069 AGCAAGAGCCAGCAAAAGCCAGGACCGTTAAAGAGCCCGCTTGTGCGTTTTCCTCA 3128
Db 2850 AGCAAGAGCCAGCAAAAGCCAGGACCGTTAAAGAGCCCGCTTGTGCGTTTTCCTCA 2791
QY 3129 TAGGTCGCGCCCTTCGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 3188

Db	2730		TAGGCTCGCCGCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGTGGCGAAA	2731
Qy	3189		CCGACAGGATATATAAGATACCAAGCGCTTTCCCTCGGAAGCTCCCTCGTGGCTCTCC	3248
Db	2730		CCGACAGGATATATAAGATACCAAGCGCTTTCCCTCGGAAGCTCCCTCGTGGCTCTCC	2671
Qy	3249		TGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGC	3308
Db	2670		TGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGC	2611
Qy	3309		GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTGGTGTAGGTCTGCTCCAAAGCT	3368
Db	2610		GCTTTCTCAAAGCTCACGCTGTAGGTATCTCAGTTTGGTGTAGGTCTGCTCCAAAGCT	2551
Qy	3369		GGGCTGTGTGCAGAAACCCCGCTTACGCCAGCCGCTGCGCTTATCCGGTAACATCG	3428
Db	2550		GGGCTGTGTGCAGAAACCCCGCTTACGCCAGCCGCTGCGCTTATCCGGTAACATCG	2491
Qy	3429		TCTTTAGTCCAAACCGGTAAAGACAGACTTATCGCCTCGCAGCAGCACTGGTAACAG	3488
Db	2490		TCTTTAGTCCAAACCGGTAAAGACAGACTTATCGCCTCGCAGCAGCACTGGTAACAG	2431
Qy	3489		GATTAGCAGCGAGGTATGTAGCGGTGTACAGAGTTCTTGAAGTGGTGGCTTAACTA	3548
Db	2430		GATTAGCAGCGAGGTATGTAGCGGTGTACAGAGTTCTTGAAGTGGTGGCTTAACTA	2371
Qy	3549		CGGCTACACTAGAAACAGTATTTGGTATCTCGGCTCTGCTGAGCCAGTTACCTCGG	3608
Db	2370		CGGCTACACTAGAAACAGTATTTGGTATCTCGGCTCTGCTGAGCCAGTTACCTCGG	2311
Qy	3609		AAAAAGAGTTGGTAGCTCTTGATCCGCCAAACAAACACCGCTGTTAGCGGTGGTTTTT	3668
Db	2310		AAAAAGAGTTGGTAGCTCTTGATCCGCCAAACAAACACCGCTGTTAGCGGTGGTTTTT	2251
Qy	3669		TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGATCTCAAGAGATCTTTGTGATCTT	3728
Db	2250		TGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGATCTCAAGAGATCTTTGTGATCTT	2191
Qy	3729		TTCTACGGGTCTGACGCTCAGTGGAAAGAAACTCAAGTTAAGGATTTTGTCTATGAG	3788
Db	2190		TTCTACGGGTCTGACGCTCAGTGGAAAGAAACTCAAGTTAAGGATTTTGTCTATGAG	2131
Qy	3789		ATTATCGTCG-----	3798
Db	2130		TTTTCGCTGCATAACCTGCTTCGGGGTCATTATAGCGATTTTTCGGTATATCCATCC	2071
Qy	3799		-----	3798
Db	2070		TTTTTCGCAGATATACAGATTTTGCAAGGGTTCGTGTAGACTTTCTTGTTGTATC	2011
Qy	3799		-----ACCAAGCGGCATCGTGCCTCCCACTCCCGACCTCTGCAAGTCCGGG	3840
Db	2010		CAACGGGTTCAGCGGGCAGGATAGGTGAAGTAGGCCCAACCGCAGCGGGTGTCTCTTC	1951
Qy	3841		GCATGGATCGCGGATAGCGCTGCTGTTTCTGTGATCCGACGATTTGCACTGCCGG	3900
Db	1950		TTCACTGTCCCTTATTCGCACTTGGCGTGTCTCAAACGGGAATCCTGCTCTGGAGGCTGG	1891
Qy	3901		TAGAACTCCGC-----GAGGTGCTCCAGCCTCAGCAGCAGCTGAACCACT	3947
Db	1890		CCGGCTACCGCGCGGTAAACAGATGAGGGCAACGGATGGCTGATGAACCAAGCCAACC	1831
Qy	3948		CGCAGGGGATCGAGCCGGGTGGCGAAGAACTCCAGATGAGATCCCGCGCTGGAG	4007
Db	1830		AGAAAGGCGAGCCCACTTATCAAGGTTGTACTGCTTCCAGACGAACGAAGCGATTGAG	1771
Qy	4008		GATCATCCAG-----	4018
Db	1770		GAAGAAGCGCGCGCGCATGAGCCTGTCCGCCCTACTGCTGGCGTCCGCCAGGCG	1711
Qy	4019		-----CGGCGTCCCGAAAAACGAT	4037


```

QY 5118 AGCCACTGCAAGCTACTGCTTCTCTTTGGCTTGGCTTTCCTTGTCCATGATAGCC 5177
Db 570 AGCCACTGCAAGCTACTGCTTCTCTTTGGCTTGGCTTTCCTTGTCCATGATAGCC 511
QY 5178 CAGTAGCTGACATTCATCCGGGTCAGACACCGTTTCTGCGACTGGCTTCTAGCTGTT 5237
Db 510 CAGTAGCTGACATTCATCCGGGTCAGACACCGTTTCTGCGACTGGCTTCTAGCTGTT 451
QY 5238 CGCTTCCTTTAGACACCTTGGCCCTCAGTGCCTTGGCGACGCTG 5283
Db 450 CGCTTCCTTTAGACACCTTGGCCCTCAGTGCCTTGGCGACGCTG 405

RESULT 11
AY228222/c
LOCUS      5695 bp      DNA      circular SYN 25-MAR-2003
DEFINITION      Shuttle vector pBC-K19mob2, complete sequence.
ACCESSION      AY228222
VERSION        AY228222.1 GI:29242905
KEYWORDS
SOURCE
ORGANISM      Shuttle vector pBC-K19mob2
                Shuttle vector pBC-K19mob2
                artificial sequences; vectors.
REFERENCE      1 (bases 1 to 5695)
AUTHORS      Kirchner,O. and Tauch,A.
TITLES      Tools for genetic engineering in the amino acid-producing bacterium
                Corynebacterium glutamicum
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 5695)
AUTHORS      Kirchner,O. and Tauch,A.
TITLES      Direct Submission
JOURNAL      Submitted (29-JAN-2003) Department of Genetics, University of
                Bielefeld, Universitaetsstrasse 25, Bielefeld D-33615, Germany
FEATURES
    source
        1..5695
            /organism="Shuttle vector pBC-K19mob2"
            /mol_type="genomic DNA"
            /db_xref="taxon:224992"
            /note="Escherichia coli-Corynebacterium glutamicum shuttle
                vector"
    gene
        400..1863
            /gene="repa"
    CDS
        400..1863
            /gene="repa"
            /note="replication protein"
            /codon_start=1
            /transl_table=11
            /product="Repa"
            /protein_id="AA066604.1"
            /db_xref="GI:29242906"
            /translation="MTLADPQDTVTASAWKLSADLPDTHPEAMRCGSRGWTADREEL
                LAHGRBSFGSKTRDFASAWIKNPOTGETQPKLYRAGSKALTRCOYVALTHAQAAV
                IVLIDVPSHQAGSKIEHVNPOYVAILLEKWARLEKAPWIGVNPPLSGKQCLINLIDPV
                YAAKQSPNRLIAITEEMRVFGADQAFSHLSRWPLHVSDDPTAYKWHGCHDRV
                DLADLMELARTWTGSKPKYIEQDFSSGRARI EAAQRATAEKALALIDASLPAL
                DASGLIDGVYRWNTNDETFARHALTVGYQLKAGERLKDAKALIDAVEAYNV
                AQAAGADGREDPLPAMDRDLTMARVRGVYAKGPVYPARRVETQSSRGKALATWGR
                RGAATSNARWADPESKVAOETRORLAEANKREMTGELLELVKLTALIDARSQSVD
                PSTRELAGELVGSERRIQQVRKALGMEAKGRPAEN"
            2505..2721
                /gene="per"
            /gene="per"
            /note="positive effector of replication"
            /codon_start=1
            /transl_table=11
            /product="per"
            /protein_id="AA066605.1"
            /db_xref="GI:29242907"
            /translation="MDLSVDVALESALDAVLDLTKVIGFFRALDITDAPASRDWASAA
                SDLETLVADLELADELARQRQEDAQ"
            3240..4034
                /gene="aph(3')-IIa"
                /note="aminoglycoside phosphotransferase"
                /codon_start=1
                /transl_table=11
                /product="Aph(3')-IIa"
                /protein_id="AA066606.1"
                /db_xref="GI:29242908"
                /translation="MIEODGLHAGSPAAVERLFGVDMAQQTIGSDAAVPELSAGGR
                    PVLVFKTDLGALNELQDEARLSWLTGVCFAVLVDVTEAGRDMLLGLVFGQDL
                    LSHLAPAEKVISIMADARMLHLDPATCFPDHQAKRIERARTMEAGLVDDDLDE
                    EHQGLAPAEFLFARKRMPDGLDVLVTHGDACLPMINVENGFSFGFIDGRLGVADRY
                    ODIALATRDIAEELGGEWADRLVLYGIAAPDSQRIAFYRLIDEPF"
                    4419..4517
                        /note="RP4-mob region"
                    5280..5282
                        /note="oriV"
BASE COUNT      1285 a 1626 c 1636 g 1148 t
ORIGIN
Query Match      39.5%; Score 2088.4; DB 12; Length 5695;
Best Local Similarity 86.7%; Pred. No. 0;
Matches 2434; Conservative 0; Mismatches 141; Indels 231; Gaps 3;
QY 2709 GACGTAATCATGCTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAAATTCACAA 2768
Db 5695 GCGGTAATCATGCTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAAATTCACAA 5636
QY 2769 CAACATACGAGCGGGAAGCATAAAGTGTAAAGCCGCGGTGCTTAATGAGTGAGCTAACT 2828
Db 5635 CAACATACGAGCGGGAAGCATAAAGTGTAAAGCCGCGGTGCTTAATGAGTGAGCTAACT 5576
QY 2829 CACATTAATTCGTTTCGCTCTACTCCCGCTTTCAGTTCGGGAAACCTGTCTGTCAGCT 2888
Db 5575 CACATTAATTCGTTTCGCTCTACTCCCGCTTTCAGTTCGGGAAACCTGTCTGTCAGCT 5516
QY 2889 GCATTAATGAATTCGGCCCAACGCGGGGAGAGCGGTTTGGTATTTGGCGCTCTTCGCG 2948
Db 5515 GCATTAATGAATTCGGCCCAACGCGGGGAGAGCGGTTTGGTATTTGGCGCTCTTCGCG 5456
QY 2949 TTCTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCA 3008
Db 5455 TTCTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCA 5396
QY 3009 CTCAAAGCGGTATACAGGTTATCCACAGATACAGGGGATACGCGAGAAACATGTG 3068
Db 5395 CTCAAAGCGGTATACAGGTTATCCACAGATACAGGGGATACGCGAGAAACATGTG 5336
QY 3069 AGCAAAAGGCCAGCAAAAGGCCAGCAACCGTAAAGAGCCGCTTGTGCGGCTTTTCCCA 3128
Db 5335 AGCAAAAGGCCAGCAAAAGGCCAGCAACCGTAAAGAGCCGCTTGTGCGGCTTTTCCCA 5276
QY 3129 TAGGTCGCGCCCTCTGACGAGCATCAAAAATTCGACGCTCAAGTCAGAGGTGGCGGAAA 3188
Db 5275 TAGGTCGCGCCCTCTGACGAGCATCAAAAATTCGACGCTCAAGTCAGAGGTGGCGGAAA 5216
QY 3189 CCGGACAGACTATTAAGATACACAGCGGTTTCCCGCTGGAAGCTCCCTCGCTCGCTTCC 3248
Db 5215 CCGGACAGACTATTAAGATACACAGCGGTTTCCCGCTGGAAGCTCCCTCGCTCGCTTCC 5156
QY 3249 TGTTCCGACCTTCCGCTTACCGGATACCTGTCTCGCCCTTTTCTCCCTTCGGGAAGCGTGGC 3308
Db 5155 TGTTCCGACCTTCCGCTTACCGGATACCTGTCTCGCCCTTTTCTCCCTTCGGGAAGCGTGGC 5096
QY 3309 GCTTTCTCATAGCTCACGCTGTAGTATCTCAGTTCGCTGAGTTCGCTCGCTCAAGCT 3368
Db 5095 GCTTTCTCATAGCTCACGCTGTAGTATCTCAGTTCGCTGAGTTCGCTCGCTCAAGCT 5036
QY 3369 GGGCTGTGTGACGAACCCCGCTTTCAGCCCGACCGCTTTCAGCCCGACCGCTTATCCGGTAACTATCG 3428
Db 5035 GGGCTGTGTGACGAACCCCGCTTTCAGCCCGACCGCTTTCAGCCCGACCGCTTATCCGGTAACTATCG 4976

```

3429 TCTTGAGTCAACCCCGTAAAGACAGACTTATGCGCACTGGCAGCAGCCACTGTAACAG 3488
Db
4975 TCTTGAGTCAACCCCGTAAAGACAGACTTATGCGCACTGGCAGCAGCCACTGTAACAG 4916
Qy
3489 GATTACAGAGCAGAGGTATGTAAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTTAACCTA 3548
Db
4915 GATTACAGAGCAGAGGTATGTAAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTTAACCTA 4856
Qy
3549 CGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGTAAGCCAGTTACCTTCGG 3608
Db
4855 CGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGTAAGCCAGTTACCTTCGG 4796
Qy
3609 AAAAAGAGTGGTAGCTCTGATCCGGCAAAACAAACCCAGCTGCTAGCGGTGGTTTTTT 3668
Db
4795 AAAAAGAGTGGTAGCTCTGATCCGGCAAAACAAACCCAGCTGCTAGCGGTGGTTTTTT 4736
Qy
3669 TGTTCGAAGCAGCAGATACCGCAGAGAAAAGATCTCAAGAGATCTCTTTGATCTT 3728
Db
4735 TGTTCGAAGCAGCAGATACCGCAGAGAAAAGATCTCAAGAGATCTCTTTGATCTT 4676
Qy
3729 TTCTACGGGTCTGAGCTCAGTGGAAACGAAACTCACTTAAGGGATTTTGGTCAATGAG 3788
Db
4675 TTCTACGGGTCTGAGCTCAGTGGAAACGAAACTCACTTAAGGGATTTTGGTCAATGAG 4616
Qy
3789 ATTATCGTCG----- 3798
Db
4615 TTTTCGCTGCATAACCTGCTTCGGGTCAATATAGCGATTTTTCGGTATATCCATCC 4556
Qy
3799 ----- 3798
Db
4555 TTTTCGCGCATATACAGGATTTTGCAGAGGTTCTGTTAGACTTTCTTCTGGTGTATC 4496
Qy
3799 -----ACCAAGCGCCATCGTGCCTCCCACTCTCTGAGTTCCGGG 3840
Db
4495 CAACGGCTCAGGGGGAGGATAGTGAAGTAGGCGCCACCGCAGCGGGTTCCTTC 4436
Qy
3841 GCATGATCGCGGATAGCGCTGCTGTTTCTGATGCCAGCGATTTGCACTGCGG 3900
Db
4435 TTCACTGTCCTTATTCGCACCTGGCGGTCTCAACGGGAATCTCTCTGCGAGGCTG 4376
Qy
3901 TAGAATCCGC-----GAGTCTGTCAGCTCAGGAGCAGCTGMAACCACT 3947
Db
4375 CCGGCTACCGCGCGGCTACAGATAGAGGCAAGCGGATGGCTGATGAACCAAGCAACC 4316
Qy
3948 CGGAGGGGATCGAGCCCGGGTGGCGGAAGAACTCCAGCATGAGATCCCGCGCTGGAG 4007
Db
4315 AGGAAGGCGAGCCACTATCAGGTGTACTGCTTCAGACGAACGAGAGCGATTGAG 4256
Qy
4008 GATCATCCAG----- 4018
Db
4255 GAAAGCGCGCGCGCGCATAGCCTGTGCGGCTACTCTGGCGGTGGCCAGGCG 4196
Qy
4019 -----CGGCTCCCGGMAACGAT 4037
Db
4195 TACAAAATCAGGGGCTGTGACTATGACGAGTCCGGAGGGCGTCCCGGAAACGAT 4136
Qy
4038 TCGAAGCCCAACTTTCATAGAAGCGCGGTGGAATCGAAATCTCGTATGCGAGTT 4097
Db
4135 TCGAAGCCCAACTTTCATAGAAGCGCGGTGGAATCGAAATCTCGTATGCGAGTT 4076
Qy
4098 GGGCGTCTGCTGCTGCTATTCGAACCCAGAGTCCCGCTCAGAGAACTCTGTCAGA 4157
Db
4075 GGGCGTCTGCTGCTGCTATTCGAACCCAGAGTCCCGCTCAGAGAACTCTGTCAGA 4016
Qy
4158 AGGCGATAGAAGCGGATGGCTGCGAATCGGGAGCGCGGATACCGTAAAGCACAGGAG 4217
Db
4015 AGGCGATAGAAGCGGATGGCTGCGAATCGGGAGCGCGGATACCGTAAAGCACAGGAG 3956
Qy
4218 CGGTACGCCCATTCGCGCGGAGCTCTTCAAGCAATATCAGGGTAGCCAACTATGTCC 4277
Db
3955 CGGTACGCCCATTCGCGCGGAGCTCTTCAAGCAATATCAGGGTAGCCAACTATGTCC 3896
Qy
4278 TGATAGCGGTCCGCCACACCCAGCGGCCACAGTCGATGAATCCAGAAAAGCGCAATT 4337

Db
3895 TGATAGCGGTCCGCCACACCCAGCGGCCACAGTCGATGAATCCAGAAAAGCGCCATT 3836
Qy
4338 TCCACCATGATATTCGGCAAGCAGGATCGCCATGGTCAAGAGATCTCTCGCGTGC 4397
Db
3835 TCCACCATGATATTCGGCAAGCAGGATCGCCATGGTCAAGAGATCTCTCGCGTGC 3776
Qy
4398 GGCATCGCGCTTGAGCCCTGGCGAAAGTTCGGCTGGCGCGAGCCCTGATGCTCTTCG 4457
Db
3775 GGCATCGCGCTTGAGCCCTGGCGAAAGTTCGGCTGGCGCGAGCCCTGATGCTCTTCG 3716
Qy
4458 TCCAGATCATCTGATGTCAGACAGCGCTTCCATCCGAGTACGTCCTCGATGCGA 4517
Db
3715 TCCAGATCATCTGATGTCAGACAGCGCTTCCATCCGAGTACGTCCTCGATGCGA 3656
Qy
4518 TGTTCGCTTGGTGGTTCGAATGGCAGAGTAGCCGATCAAGCGTATGACGCGCCGCAAT 4577
Db
3655 TGTTCGCTTGGTGGTTCGAATGGCAGAGTAGCCGATCAAGCGTATGACGCGCCGCAAT 3596
Qy
4578 GCATCAGCCATGATGATATCTTCTCGGAGAGGAGAGTGATGATGACAGGAGATCTTCG 4637
Db
3595 GCATCAGCCATGATGATATCTTCTCGGAGAGGAGAGTGATGATGACAGGAGATCTTCG 3536
Qy
4638 CCGGCACTTCGCGCAATAGCAGCAGTCCCTTCCGCTTCAGTGACAACTGACAGCA 4697
Db
3535 CCGGCACTTCGCGCAATAGCAGCAGTCCCTTCCGCTTCAGTGACAACTGACAGCA 3476
Qy
4698 GTTCGCGAAGAAAGCGCGTCTGCGCAGCAGATAGCGCGCTGCTCTCTCAGT 4757
Db
3475 GTTCGCGAAGAAAGCGCGTCTGCGCAGCAGCAGATAGCGCGCTGCTCTCTCAGT 3416
Qy
4758 TCATTCAGGGCACCGGACAGTGGTCTTGAACAAAGAACCGGGCGCCCTGCGCTGAC 4817
Db
3415 TCATTCAGGGCACCGGACAGTGGTCTTGAACAAAGAACCGGGCGCCCTGCGCTGAC 3356
Qy
4818 AGCGGAAACACCGCGCATCAGAGCAGCGATGTCTGTGTGCCAGTCAATGCGCAAT 4877
Db
3355 AGCGGAAACACCGCGCATCAGAGCAGCGATGTCTGTGTGCCAGTCAATGCGCAAT 3296
Qy
4878 AGCGGAAACACCGCGCATCAGAGCAGCGATGTCTGTGTGCCAGTCAATGCGCAAT 4937
Db
3295 AGCGGAAACACCGCGCATCAGAGCAGCGATGTCTGTGTGCCAGTCAATGCGCAAT 3236
Qy
4938 AAGCATCTCTCATCTCTCTCTGATCAGATCTTGTATCCCTGCGCATCAGATCTTGGC 4997
Db
3235 AAGCATCTCTCATCTCTCTGATCAGATCTTGTATCCCTGCGCATCAGATCTTGGC 3176
Qy
4998 GGCAGAAAGCCATCCAGTTTACTTTGAGGGCTTCCAAACCTTACAGAGGGCGCCCA 5057
Db
3175 GGCAGAAAGCCATCCAGTTTACTTTGAGGGCTTCCAAACCTTACAGAGGGCGCCCA 3116
Qy
5058 GCTGGCAATTCGGTTTCGCTTGTGTCATATAAACCGCCAGTCTAGCTATCGCCATGTA 5117
Db
3115 GCTGGCAATTCGGTTTCGCTTGTGTCATATAAACCGCCAGTCTAGCTATCGCCATGTA 3056
Qy
5118 AGCCACTGCAAGCTACCTGCTTCTTGTGGCTTGGCTTTCCTTTCCTTTCAGATAGCC 5177
Db
3055 AGCCACTGCAAGCTACCTGCTTCTTGTGGCTTGGCTTTCCTTTCCTTTCAGATAGCC 2996
Qy
5178 CAGTAGTGATCATTCAGGGGTGAGCAGCGTTTCTGCGAGTGGCTTCTTACGTTTC 5237
Db
2995 CAGTAGTGATCATTCAGGGGTGAGCAGCGTTTCTGCGAGTGGCTTCTTACGTTTC 2936
Qy
5238 CGCTTCCTTTAGCAGCCCTTTCGCCCTGAGTGTCTTCGCGAGCGTG 5283
Db
2935 CGCTTCCTTTAGCAGCCCTTTCGCCCTGAGTGTCTTCGCGAGCGTG 2890

RESULT 12
AY219861/c
LOCUS
DEFINITION Shuttle vector pEC-K19MECA2, complete sequence.
ACCESSION AY219861

AY219861 5869 bp DNA circular SYN 23-MAR-2003
Shuttle vector pEC-K19MECA2, complete sequence.

```
VERSION      AY219861.1  GI:29164983
KEYWORDS
SOURCE       Shuttle vector pEC-K19MECA2
ORGANISM     Shuttle vector pEC-K19MECA2
REFERENCE    1 (bases 1 to 5869)
              artificial sequences; vectors.
AUTHORS      Kirchner,O. and Tauch,A.
TITLE        Tools for genetic engineering in the amino acid-producing bacterium
              Corynebacterium glutamicum
JOURNAL      Unpublished
REFERENCE    2. (bases 1 to 5869)
              Kirchner,O. and Tauch,A.
AUTHORS      Direct Submission
TITLE        Submitted (15-JAN-2003) Department of Genetics, University of
              Bielefeld, Universitaetsstrasse 25, Bielefeld D-33615, Germany
JOURNAL
FEATURES     Location/Qualifiers
              1..5869
                 /organism="Shuttle vector pEC-K19MECA2"
                 /mol_type="genomic DNA"
                 /db_xref="taxon:224630"
                 /note="Escherichia coli-Corynebacterium glutamicum shuttle
              vector"
              574..2037
                 /gene="repA"
              574..2037
                 /gene="repA"
                 /note="replication protein"
                 /codon_start=1
                 /transl_table=11
                 /product="repA"
                 /protein_id="AA065228.1"
                 /db_xref="GI:29164984"
                 /translation="MTLADPQDTVTASAKLSADLFDTHPEAMRCGSGRWTAEDRREL
              LAHLGRSFGSKTRDPYSAWIKNPDTGETPKLYRAGSKALTRCOYVALTHAQAAY
              IVLDDVPSHQAGCKIEHVNPOVVAILEKWARLEKAPANTGUNPLSKQLWLIDPV
              YAAAGKTPSNWLLAATTEMTKRVFGADQAFSLRSLRWPLHVSDDPTAYKNWQCHDRV
              DRALDMEIARTWIGSKPKYIEQDFSSGRARIEAQAATAEAKALAILDASLEFAL
              DASGLLDIVRVLTNPERARDETAFRHALTVGYQLKAAGERLKAKIIDAYEVAYNV
              AOVAGDGRPEPDMARDRLTMARVGRYVAKQPVVPARRVETOSRGRKALATMGR
              RGAATSNARWADPESKYAQETQRLAEANKRREMTGELLELRVKTALDARSQSVD
              PSTRELAGELVGBERRIQVQRKALGMEAKRGRPAEN"
              2680..2895
                 /gene="per"
              2680..2895
                 /gene="per"
                 /note="positive effector of replication"
                 /codon_start=1
                 /transl_table=11
                 /product="per"
                 /protein_id="AA065229.1"
                 /db_xref="GI:29164985"
                 /translation="MOLSDVALESALDAVADLKTIVGFPRALDITTDAPASRDWASAA
              SDLETLVADLEELADELRARQREDAQ"
              3414..4208
                 /gene="aph(3')-IIa"
              3414..4208
                 /gene="aph(3')-IIa"
                 /note="aminoglycoside phosphotransferase"
                 /codon_start=1
                 /transl_table=11
                 /product="Aph(3')-IIa"
                 /protein_id="AA065230.1"
                 /db_xref="GI:29164986"
                 /translation="WISQGLHAGSPAANWVERFGYDMAQQTIGCSDAAVFLSAGQR
              FVLVFTKDLSEALNEQDEARLSWLAATGVPQCAVLVDVYVEAGRMDLLGVPGQDL
              LSHLPAEYVSMADARLHLDLPATCFDFQAKHRIERFARTRMEAGLIVPGQDLDE
              EHOGLAPAEFLAKRMPOGDELVTGHDACLPNTVNGRPSGFDICGRIGVADRY
              QDIALATRDIAEELGGEWADRFVLVGLIAPDQSRIAFVRLIDEFF"
              4593..4691
                 /misc_feature
                 /note="RP4-mob region"
              5454..5456
                 /rep_origin
                 /note="oriV"
              BASE COUNT      1325 a 1669 c 1683 g 1192 t
```

ORIGIN

```
Query Match      39.5%; Score 2088.4; DB 12; Length 5869;
Best Local Similarity 86.7%; Pred. No. 0;
Matches 2434; Conservative 0; Mismatches 141; Indels 231; Gaps 3;

QY 2709 GACGTAATCATGGTCATAGCTGTTCCCTGCTGGAATTTGTTATCCGCTCACAATTCACA 2768
DB 5869 GCGGTAATCATGGTCATAGCTGTTCCCTGCTGGAATTTGTTATCCGCTCACAATTCACA 5810

QY 2769 CAACATACGAGCCGGAAGCATAAAGTGTAAAGCCCTGGGGTGCCTTAATGAGTGAGCTAACT 2828
DB 5809 CAACATACGAGCCGGAAGCATAAAGTGTAAAGCCCTGGGGTGCCTTAATGAGTGAGCTAACT 5750

QY 2829 CACATTAATTCGGTTGCGCTCACTGCCCGCTTCCAGTCGCGGAACCTGTGTCGCCAGCT 2888
DB 5749 CACATTAATTCGGTTGCGCTCACTGCCCGCTTCCAGTCGCGGAACCTGTGTCGCCAGCT 5690

QY 2889 GCATTAATGAATCGGCCCAACCGCGGGGAGAGGGGGTTTGGGTATTTGGCGCGCTCTTCGCG 2948
DB 5689 GCATTAATGAATCGGCCCAACCGCGGGGAGAGGGGGTTTGGGTATTTGGCGCGCTCTTCGCG 5630

QY 2949 TTCCTCGCTCACTGACTCGCTGCGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTC 3008
DB 5629 TTCCTCGCTCACTGACTCGCTGCGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGTTC 5570

QY 3009 CTCGAAAGCGGGTAAATACCGTTATCCACAGAAATCAGGGGATTAACCGCAGGAAAGACATGTG 3068
DB 5569 CTCGAAAGCGGGTAAATACCGTTATCCACAGAAATCAGGGGATTAACCGCAGGAAAGACATGTG 5510

QY 3069 AGCAAAAGCGCCAGCAAAAGGCGCAGAACCGTAAAGAGCCCGTGTGTGGCGTTTTCCTCA 3128
DB 5509 AGCAAAAGCGCCAGCAAAAGGCGCAGAACCGTAAAGAGCCCGTGTGTGGCGTTTTCCTCA 5450

QY 3129 TAGGTCGCGCGCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGGAAA 3188
DB 5449 TAGGTCGCGCGCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGGAAA 5390

QY 3189 CCCGACAGAGCTATAAAGATACCAAGCGGCTTCCCGCTCGGAAGCTCCCTTCGTCGCGCTCTCC 3248
DB 5389 CCCGACAGAGCTATAAAGATACCAAGCGGCTTCCCGCTCGGAAGCTCCCTTCGTCGCGCTCTCC 5330

QY 3249 TGTTCCGACCTTCGCGCTTACGGATACCTGTCGCGCTTTCCTCCCTCGGGAAGCGTGGC 3308
DB 5329 TGTTCCGACCTTCGCGCTTACGGATACCTGTCGCGCTTTCCTCCCTCGGGAAGCGTGGC 5270

QY 3309 GCTTTCTCATAGCTCAGCGCTGTAGGTATCTCAGTTTCGGTGTAGTTCGTTCCGCTCAAGCT 3368
DB 5269 GCTTTCTCATAGCTCAGCGCTGTAGGTATCTCAGTTTCGGTGTAGTTCGTTCCGCTCAAGCT 5210

QY 3369 GGGCTGTGTGACGAAACCCCGCTTCAGCCCGACCGCTTCGCGCTTATCCCGTAACTATCG 3428
DB 5209 GGGCTGTGTGACGAAACCCCGCTTCAGCCCGACCGCTTCGCGCTTATCCCGTAACTATCG 5150

QY 3429 TCTTAGTTCCAACCGGTAAAGACACGACTTATCCGCTGCGCAGCCACTGGTAAACAG 3488
DB 5149 TCTTAGTTCCAACCGGTAAAGACACGACTTATCCGCTGCGCAGCCACTGGTAAACAG 5090

QY 3489 GATTAGCAGAGCGAGGTATGTAGGCGGTCTACAGAGTTCCTTGAAGTGGTGGCTTAACCTA 3548
DB 5089 GATTAGCAGAGCGAGGTATGTAGGCGGTCTACAGAGTTCCTTGAAGTGGTGGCTTAACCTA 5030

QY 3549 CGGCTACACTAGAGAACAGTATTTGGTATTCGCGCTCTGCTGAAGCCAGTACTCTTCGG 3608
DB 5029 CGGCTACACTAGAGAACAGTATTTGGTATTCGCGCTCTGCTGAAGCCAGTACTCTTCGG 4970

QY 3609 AAAAAGAGTTCGTAGCTCTTCATCCGCAAAACCAACCCGCTGGTAGCGGTGGTGTCTTT 3668
DB 4969 AAAAAGAGTTCGTAGCTCTTCATCCGCAAAACCAACCCGCTGGTAGCGGTGGTGTCTTT 4910

QY 3669 TGTTTGAAGCAGCAGATTACGCGCAGAAAAAGGATCTCAAGAAAGATCTTTGATCTT 3728
DB 4909 TGTTTGAAGCAGCAGATTACGCGCAGAAAAAGGATCTCAAGAAAGATCTTTGATCTT 4850
```

3729 TTCTACGGGCTGACGCTCAGTGGAAACGAAAACTCAGTTAAGGATTTTGGTCAATGAG 3788
Db |||||
4849 TTCTACGGGCTGACGCTCAGTGGAAACGAAAACTCAGTTAAGGATTTTGGTCAATGAG 4790
QY |||||
3789 ATTATGCTGC----- 3798
Db |||||
4789 TTTTCCGCTGCATAACCCCTGCTTCGGGGTCAATTATAGCGATTTTTCGGTATATCCATCC 4730
QY |||||
3799 ----- 3798
Db |||||
4729 TTTTTCGCAGATATACAGGATTTTGGCAAGGGTTCGTGTAGACTTTCTTGGTGTATC 4670
QY |||||
3799 -----ACAAAGCGCCATCGCTCTCCCACTCTCTGAGTTTCCTGAGTTTCGGG 3840
Db |||||
4669 CAAAGCGCTCAGCGGGCAGGATAGTGAAGTAGCCCAACCGGAGCGGGTTCCTTC 4610
QY |||||
3841 GCATGATGCGGGATAGCGGCTGCTGTTTCTGATGCGGATTTGCACTGCGG 3900
Db |||||
4609 TTCACTGCTCTTATTCGACCTGGCGGTGCTCAACGGGAATCTCTGCTCTCGAGGCTGG 4550
QY |||||
3901 TAGAACTCGC-----GAGGTCTCTCAGCCTCAGGCGAGCTGAACCAACT 3947
Db |||||
4549 CCGGTACCGCGCGCTAACAGTAGGCAAGCGATGGGTGATGAACCAAGCAACC 4490
QY |||||
3948 CCGGAGGGATCAGCGCGGGTGGCGAAGAACTCCAGCATGATCCCCCGCTGGAG 4007
Db |||||
4489 AGGAAGGCGAGCCCACTATCAAGGTGACTGCTTCCAGACGAACGAAGAGCGATTGAG 4430
QY |||||
4008 GATCATCCAGC----- 4018
Db |||||
4429 GAAAGGCGCGCGCGCATGAGCCTGTGCGCTACCTGCTGGCGCTCGGCCAGGC 4370
QY |||||
4019 -----CGGCTCCCGGAAAAAGAT 4037
Db |||||
4369 TACAATAACAGCGGCTGCTGACTATGACACGCTCGCGAGGGCTCCCGGAAAAAGAT 4310
QY |||||
4038 TCGAAGCCCACTTTATAGAAGCGCGGTGGAATCGAAATCTCTGATGCGAGTT 4097
Db |||||
4309 TCGAAGCCCACTTTATAGAAGCGCGGTGGAATCGAAATCTCTGATGCGAGTT 4250
QY |||||
4098 GGGCTGCTGCTGCTGCTTCAAGCCAGATGCTCGGTTCAGAAATCTGTCAGAA 4157
Db |||||
4249 GGGCTGCTGCTGCTGCTTCAAGCCAGATGCTCGGTTCAGAAATCTGTCAGAA 4190
QY |||||
4158 AGCGATAGAGCGATGCTGCGAATCGGAGCGCGATACCGTAAGCACAGGAAG 4217
Db |||||
4189 AGCGATAGAGCGATGCTGCGAATCGGAGCGCGATACCGTAAGCACAGGAAG 4130
QY |||||
4218 CGGTACGCCATTCGCGCCAAAGCTCTTCAGCAATATCACGGGTAGCCAAAGCTATGTCC 4277
Db |||||
4129 CGGTACGCCATTCGCGCCAAAGCTCTTCAGCAATATCACGGGTAGCCAAAGCTATGTCC 4070
QY |||||
4278 TGTAGCGGTCCGCAACAGCGCGGACAGTGTATGATATCCAGAAAGCGGCATTT 4337
Db |||||
4069 TGATAGCGGTCCGCAACAGCGCGGACAGTGTATGATATCCAGAAAGCGGCATTT 4010
QY |||||
4338 TCACCATGATATTCGCGCAAGAGGATCGCCATGGTCAACACAGATCTTCGCGCTCG 4397
Db |||||
4009 TCACCATGATATTCGCGCAAGAGGATCGCCATGGTCAACACAGATCTTCGCGCTCG 3950
QY |||||
4398 GGCATCGCGCTTGAAGCTGGGAAACAGTTTCGCTGGCGGAGCGCCCTGATGCTTCG 4457
Db |||||
3949 GGCATCGCGCTTGAAGCTGGGAAACAGTTTCGCTGGCGGAGCGCCCTGATGCTTCG 3890
QY |||||
4458 TCAGATATCTCTGATTCGAACAGCGGCTTCATCCAGTACGTGCTGCTGATGCGA 4517
Db |||||
3889 TCAGATATCTCTGATTCGAACAGCGGCTTCATCCAGTACGTGCTGCTGATGCGA 3830
QY |||||
4518 TGTTCGCTTGGTTCGAATGGCGAGGTAGCGGATCAAGCGATGACGCGCGCGCAT 4577
Db |||||
3829 TGTTCGCTTGGTTCGAATGGCGAGGTAGCGGATCAAGCGATGACGCGCGCGCAT 3770

QY 4578 GCATCAGCCATGATGATATCTTCTCGCAGGAGCAAGTGTAGATGACAGGAGATCTCTGC 4637
Db |||||
3769 GCATCAGCCATGATGATATCTTCTCGCAGGAGCAAGTGTAGATGACAGGAGATCTCTGC 3710
QY |||||
4638 CCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCCGCTTCAGTCAACGTCGAGCACA 4697
Db |||||
3709 CCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCCGCTTCAGTCAACGTCGAGCACA 3650
QY |||||
4698 GTTGGCAAGGAACCGCTGCTGCGCAGCAGCAGTACGCGGCTGCTCTGCTGAGT 4757
Db |||||
3649 GCTGCGCAAGGAACCGCTGCTGCGCAGCAGCAGTACGCGGCTGCTCTGCTGAGT 3590
QY |||||
4758 TCATTAGGGCAGCGCAGGTCGCTCTGTGCAAAAAGAACCGGGGCGCTCGCTGAC 4817
Db |||||
3589 TCATTAGGGCAGCGCAGGTCGCTCTGTGCAAAAAGAACCGGGGCGCTCGCTGAC 3530
QY |||||
4818 AGCGGAAACACCGCGCATCAGAGCAGCGATGCTGTGTGCGCCAGTCATAGCCGAT 4877
Db |||||
3529 AGCGGAAACACCGCGCATCAGAGCAGCGATGCTGTGTGCGCCAGTCATAGCCGAT 3470
QY |||||
4878 AGCTCTCCACCGAGCGCGGAGAACCTGGTGCATTCATCTTGTTCATCATGGA 4937
Db |||||
3469 AGCTCTCCACCGAGCGCGGAGAACCTGGTGCATTCATCTTGTTCATCATGGA 3410
QY |||||
4938 AACGATCTCTCTCTCTGATCAGATCTTGTGATCCCTGCGCATCAGATCTTGGC 4997
Db |||||
3409 AACGATCTCTCTCTCTGATCAGATCTTGTGATCCCTGCGCATCAGATCTTGGC 3350
QY |||||
4998 GGCAGAAAGCCATCCAGTTTCTTTCAGGCTTCCCAACCTTACAGAGGCGGCCCA 5057
Db |||||
3349 GGCAGAAAGCCATCCAGTTTCTTTCAGGCTTCCCAACCTTACAGAGGCGGCCCA 3290
QY |||||
5058 GCTGGCAATTCGGTTCGCTGCTGCTCCATAAAACCGCCAGTCTAGCTATCGCCATGA 5117
Db |||||
3289 GCTGGCAATTCGGTTCGCTGCTGCTCCATAAAACCGCCAGTCTAGCTATCGCCATGA 3230
QY |||||
5118 AGCCCACTGCAAGTACCTGCTTCTTTCGCTTTCGCTTTCCTTTCCTGTCAGATAGCC 5177
Db |||||
3229 AGCCCACTGCAAGTACCTGCTTCTTTCGCTTTCGCTTTCCTTTCCTGTCAGATAGCC 3170
QY |||||
5178 CAGTACGATGATTCATCGGGGTGAGCACCTTTCGCGAGTGGCTTTCAGCTGTTTC 5237
Db |||||
3169 CAGTACGATGATTCATTCGGGTGAGCACCTTTCGCGAGTGGCTTTCAGCTGTTTC 3110
QY |||||
5238 CGCTTCTTTAGCAGCCCTTTCGCGCTGAGTGTCTGCGCAGCGTG 5283
Db |||||
3109 CGCTTCTTTAGCAGCCCTTTCGCGCTGAGTGTCTGCGCAGCGTG 3064

RESULT 13
AF445080/cLOCUS 5695 bp DNA circular SYN 18-SEP-2002
DEFINITION Cloning vector pEC-K18mob2, complete sequence.

ACCESSION AF445080

VERSION AF445080.1 GI:17386067

KEYWORDS

SOURCE ORGANISM

Cloning vector pEC-K18mob2

Cloning vector pEC-K18mob2

artificial sequences; vectors.

1 (bases 1 to 5695)

Tauch, A., Kirchner, O., Löffler, B., Gotker, S., Puhler, A. and

Kalinowski, O.

Efficient Electroporation of Corynebacterium diphtheriae with
a Mini-Replicon Derived from the Corynebacterium glutamicum Plasmid

pgal

Curr. Microbiol. 45 (5), 362-367 (2002)

12232668

2 (bases 1 to 5695)

Tauch, A.

Direct Submission

Submitted (08-NOV-2001) Department of Genetics, University of

Bielefeld, Universitätsstrasse 25, Bielefeld D-33615, Germany

Location/Qualifiers

FEATURES

```

source
1. .5695
/organism="Cloning vector pBC-K18mob2"
/mol_type="genomic DNA"
/db_xref="taxon:179875"
/lab_host="Corynebacterium glutamicum"
184. .1647
/gene="rep"
184. .1647
/gene="rep"
/notes="rep"
/codon_start=1
/transl_table=11
/product="replication protein"
/db_xref="GI:1738580.1"
/translation="MTLADPQDTVTASAKLSADLFQTHPEAMRCGRGWTAEDEREL
LAHLGRESFGSKTRDFASAWIKNPDTGETQPKLYRAGSKALTRQVVALTHAOHAHV
IVLDDVPSHOAGGKIEHVNPOVVALEKWARLEKAPAWIGVNPGLSKCOLIWLIDPV
YAAAGTSPNRLLAATTEMTIRVFGDAQAFSHLSRPLHVSDDPTAYKWHCHQHDV
DRLADLSIARTMTGSQPKKYIIGQFSSGRSLREAAQRATPAEKALAILDASLPSAL
DASGLIDGVRVLTNPERARDETAFRHALTVGYOLKAAGERLNDKAIIDAYEVAYNV
AQAVGAGREPDLPAMRDLTMARVRGYVAKQPVVPARVETQSSRGKALATMGR
RGATSNARNRWDPSKYAQETRORLAENKREMTGELLELRVKTALILDARSOSVAD
PSTRLAGELGVSESRRIQQVRKALGMEAKRGRPRAEN"
2290. .2505
/gene="per"
2290. .2505
/gene="per"
/notes="per"
/codon_start=1
/transl_table=11
/product="positive effector of replication"
/protein_id="AAL38581.1"
/db_xref="GI:17386069"
/translation="MDLSDVALLSDALDAVDLKTIVIGFFRALDTDPAPSRDWASAA
SDLETLVADLESLADELARQRQEDAQ"
3024. .3818
/gene="aph(3')-IIa"
3024. .3818
/gene="aph(3')-IIa"
/notes="aph(3')-IIa"
/codon_start=1
/transl_table=11
/product="kanamycin resistance protein"
/protein_id="AAL38582.1"
/db_xref="GI:17386070"
/translation="MIEQDGLHAGSPAANVERLFGYDWAQQTIGCSDAAVFRLSAQR
PVLVKTLDUSGALNELQDEARLSWLTATGVPAAVLDVVTAEGRDWLLLEVEFGQDL
LSHLAPAEKVSIMADARRLHTLDPATCPFDHOKHRIERARTMEAGLVQDLDLDE
EHQGLAPAEFLARLPDGEDLVVTHGDACLINIMVENGRFSGFDICRGLGVADRY
ODIATLRDIAEELGGEWADRELVLVYIAAPDSQRIAFYRLIDBFF"
4203. .4301
/misc_feature
/notes="R24 mob region"
5064. .5066
/rep_origin
/notes="oriV of Escherichia coli"
BASE COUNT 1285 a 1624 c 1638 g 1148 t
ORIGIN

Query Match 39.5%; Score 2088; DB 12; Length 5695;
Best Local Similarity 86.8%; Pred. No. 0;
Matches 2433; Conservative 0; Mismatches 140; Indels 231; Gaps 3;

QY 2711 CGTAATCATGGTCATAGCTGTTTCCTGTGTGGAATTTGTTATCCGCTCAAAATCCACACA 2770
Db 5477 CGTAATCATGGTCATAGCTGTTTCCTGTGTGGAATTTGTTATCCGCTCAAAATCCACACA 5418
QY 2771 ACATACGAGCCGGAAGCATAAAGTGAAGCCCTGGGGTGCCTAAAGAGTGAGTAACTCA 2830
Db 5417 ACATACGAGCCGGAAGCATAAAGTGAAGCCCTGGGGTGCCTAAAGAGTGAGTAACTCA 5358
QY 2831 CATTAATTCGTTGCGGTCACCTCCCGCTTTTCCAGTCGGGAAACCTGTCTGTCGCAAGTGC 2890
Db 5357 CATTAATTCGTTGCGGTCACCTCCCGCTTTTCCAGTCGGGAAACCTGTCTGTCGCAAGTGC 5298

```

QY 3843 ATGATCGCGGATAGCGCTGCTGTTCTCTGTAGCGCGACGATTTGCACTGCGGTA 3902
Db 4217 CACTGTCTCTTATTCGACACTGGCGGTGCTCAACGGGAATCTCTGCTCTGGAGCTGGCC 4158
QY 3903 GAATCTCCGC-----GAGGTGCTCCAGCTTCAGGACGAGCTGAACCAACTCG 3949
Db 4157 GGCTACCGCGCGGCTAAACAGATGAGGGCAAGCGATGGCTGATGAACCAACCAACCAAG 4098
QY 3950 CGAGGGGATCGAGCCCGGGTGGCGGAAGAACTCCAGCATGAGATCCCGCGCTGGAGGA 4009
Db 4097 GAAGGGAGCCCACTATCAAGGTGTATCGCTTCCAGACGAAGAGCGAATTCAGGA 4038
QY 4010 TCATCCAGC----- 4018
Db 4037 AAAGCGCGCGCGCGCATGAGCTGTGGCGCTACTGTGGCGCTGCGCGCCAGGCGTA 3978
QY 4019 -----CGGCTCCCGGAACGATTC 4039
Db 3977 CAAATACAGGGCGTGTGGACTATGAGCAGCTCCGGAGGGCGTCCCGGAACGATTC 3918
QY 4040 CGAAGCCCAACCTTTTCATAGAAGCGCGGTGGAATCGAAATCTCTGTGATGGCAGGTGG 4099
Db 3917 CGAAGCCCAACCTTTTCATAGAAGCGCGGTGGAATCGAAATCTCTGTGATGGCAGGTGG 3858
QY 4100 CGTGCCTTGTGCTGCTATTTGGAACCCAGAGTCCGCTCAGNAGAACTCGTCAAGAG 4159
Db 3857 GCCTGCTTGTGCTGCTATTTGGAACCCAGAGTCCGCTCAGNAGAACTCGTCAAGAG 3798
QY 4160 CGATAGAAGCGGATGCGCTCGAATCGGAGCGCGCATACCGTAAAGACGAGGAGCG 4219
Db 3797 CGATAGAAGCGGATGCGCTCGAATCGGAGCGCGCATACCGTAAAGACGAGGAGCG 3738
QY 4220 GTCAGCCCATTCGCGCCCAAGCTCTTCAGCAATATCACGGGTAGCCAAACGCTATGCTG 4279
Db 3737 GTCAGCCCATTCGCGCCCAAGCTCTTCAGCAATATCACGGGTAGCCAAACGCTATGCTG 3678
QY 4280 ATAGGGTCCGACACACCGCGCCACAGTCCAGTATCCAGAAAGCGGCGCATTTTC 4339
Db 3677 ATAGGGTCCGCGCACACCGCGCGCCACAGTCCAGTATCCAGAAAGCGGCGCATTTTC 3618
QY 4340 CACCATGATATTCGCAAGCAGGCGATCGCCATGGTCAACGAGAGATCTTCGCGTGGG 4399
Db 3617 CACCATGATATTCGCAAGCAGGCGATCGCCATGGTCAACGAGAGATCTTCGCGTGGG 3558
QY 4400 CATGCGCGCTGAGCTGGGACAGTTCGCTGGCGGAGCGCCCTGATGCTCTTCGTC 4459
Db 3557 CATGCGCGCTGAGCTGGGACAGTTCGCTGGCGGAGCGCCCTGATGCTCTTCGTC 3498
QY 4460 CAGATCATCTGATCGAAGACCGGCTTCATCCGAGTACGTCTGCTCGATGCGATG 4519
Db 3497 CAGATCATCTGATCGAAGACCGGCTTCATCCGAGTACGTCTGCTCGATGCGATG 3438
QY 4520 TTTCGCTTGTGCTGATGGCAGGTAGCCGATCAACGCTATGACGCGCGCATTCG 4579
Db 3437 TTTCGCTTGTGCTGATGGCAGGTAGCCGATCAACGCTATGACGCGCGCATTCG 3378
QY 4580 ATCAGCCATGATGGATCTTTCTCGCGAGGAGCAAGGTGAGATGACAGAGATCTTCGCC 4639
Db 3377 ATCAGCCATGATGGATCTTTCTCGCGAGGAGCAAGGTGAGATGACAGAGATCTTCGCC 3318
QY 4640 CGGCACTTCGCCCAATAGAGCGAGTCCCTCCGCTTCAGTGACAACTGAGGACAGC 4699
Db 3317 CGGCACTTCGCCCAATAGAGCGAGTCCCTCCGCTTCAGTGACAACTGAGGACAGC 3258
QY 4700 TCGCAAGGAACCGCGCTGCTGGCGAGCCACATAGCGCGTCTGCTGCTGCTGAGTTC 4759
Db 3257 TCGCAAGGAACCGCGCTGCTGGCGAGCCACATAGCGCGTCTGCTGCTGAGTTC 3198
QY 4760 ATTACGGCACCGGACAGGTCTGTGCAAAAAGAACCGGGCGCCCTGCTGCTGACAG 4819
Db 3197 ATTACGGCACCGGACAGGTCTGTGCAAAAAGAACCGGGCGCCCTGCTGCTGACAG 3138
QY 4820 CCGGAACACGGCGCATCAGAGCAGCGGATTCCTGTTGTGCCAGTCTATAGCCGGAATAG 4879

Db 3137 CCGGAACACGGCGCATCAGAGCAGCGGATTCCTGTTGTGCCAGTCATAGCCGAATAG 3078
QY 4880 CCTCTCCACCCCAAGCGCGGAGAACCTGCGTGAATCCATCTTTGTTCAATCATGCGAAA 4939
Db 3077 CCTCTCCACCCCAAGCGCGGAGAACCTGCGTGAATCCATCTTTGTTCAATCATGCGAAA 3018
QY 4940 CGATCCTCATCTGCTCTTCATCAGATCTTGATCTCCCTCGCCCATCAGATCTTCGGCG 4999
Db 3017 CGATCCTCATCTGCTCTTCATCAGATCTTGATCTCCCTCGCCCATCAGATCTTCGGCG 2958
QY 5000 CAAGAAAGCCATCCAGTTTACTTTGCGAGGCTCCCAACCTTACAGAGGCGCGCCACG 5059
Db 2957 CAAGAAAGCCATCCAGTTTACTTTGCGAGGCTCCCAACCTTACAGAGGCGCGCCACG 2898
QY 5060 TGGCAATTCGCGTTCGCTTGCTGCTCAATAAACCCCGCAGTCTAGCTATCCGCAATGAAG 5119
Db 2897 TGGCAATTCGCGTTCGCTTGCTGCTCAATAAACCCCGCAGTCTAGCTATCCGCAATGAAG 2838
QY 5120 CCACCTGCAAGTACTCTCTTTGCGCTTGGCTTTCCCTTGCAGATAGCCCA 5179
Db 2837 CCACCTGCAAGTACTCTCTTTGCGCTTGGCTTTCCCTTGCAGATAGCCCA 2778
QY 5180 GTAGCTGCAATTCATCCGGGCTCAGCACCGTTTCTGCGGACTGGCTTTCTACGTTCG 5239
Db 2777 GTAGCTGCAATTCATCCGGGCTCAGCACCGTTTCTGCGGACTGGCTTTCTACGTTCG 2718
QY 5240 CTTCTTTAGCAGCCCTTGGCGCTGAGTGTCTGGCGCAGCGTG 5283
Db 2717 CTTCTTTAGCAGCCCTTGGCGCTGAGTGTCTGGCGCAGCGTG 2674

RESULT 14

AY222814/C

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

AY222814 3035 bp DNA circular SYN 25-MAR-2003
Cloning vector pK18mob2, complete sequence.
AY222814
AY222814.1 GI:29242887
Cloning vector pK18mob2
Cloning vector pK18mob2
artificial sequences; vectors.
1 (bases 1 to 3035)
Kirchner, O. and Tauch, A.
Tools for genetic engineering in the amino acid-producing bacterium
Corynebacterium glutamicum
Unpublished
2 (bases 1 to 3035)
Kirchner, O. and Tauch, A.
Direct Submission
Submitted (22-JAN-2003) Department of Genetics, University of
Bielefeld, Universitaetstrasse 25, Bielefeld D-33615, Germany

Location/Qualifiers
1. 3035
/organism="Cloning vector pK18mob2"
/mol_type="genomic DNA"
/db_xref="taxon:224984"
/lab_host="Escherichia coli"
584. .1378
/gene="aph(3')-IIa"
584. .1378
/gene="aph(3')-IIa"
/note="aminoglycoside phosphotransferase"
/codon_start=1
/transl_table=11
/product="Aph(3')-IIa"
/protein_id="AAO66594.1"
/db_xref="GI:29242887"
translation="MTEODLHAGSPAAVVERLFGVDVAQQTIGSDAAVFLSAQGR
PVLVYTDLSGALNELQDAARLSWLTGVCAALVDVWTEAGRDWLLGVPGDL
LSHLAPAKSVIMADMRKRLHLPATCPFHQAKRIRERATREAGLVQDDLE
EQGLAPAEFLFARLMPDGEDLVVTHGDACLPMVWENRFSFIDCGRLGVADRY
QDIALATRDIAELGGEWADRFLVLYGIAAPDSQRIAFYRLUDEFF"

gene

CDS

misc_feature		1763..1861			
/note="BP4-mob region"					
rep_origin		2624..2626			
/note="oriV"					
BASE COUNT	665 a	841 c	874 g	655 t	
ORIGIN					
Query Match 39.3%; Score 2078; DB 12; Length 3035;					
Best Local Similarity 86.7%; Pred. No. 0;					
Matches 2423; Conservative 0; Mismatches 140; Indels 231; Gaps 3;					
QY	2721	GTATAGCTGTTTCCCTGCTGTAATAATGTTATCCGCTCACAATTCACACATATACGAGC	2780		
Db	3027	GTATAGCTGTTTCCCTGCTGTAATAATGTTATCCGCTCACAATTCACACATATACGAGC	2968		
QY	2781	CGGAAGCATAAAGTGAAGAGCTGGGGTGCCTAATGAGTGAGTAACTCACATTAATTCG	2840		
Db	2967	CGGAAGCATAAAGTGAAGAGCTGGGGTGCCTAATGAGTGAGTAACTCACATTAATTCG	2908		
QY	2841	GTTCGCTCAGTCGCGCTTTCAGTCGGGAAACCTGTGTCGACGCTGCATTAATGAAT	2900		
Db	2907	GTTCGCTCAGTCGCGCTTTCAGTCGGGAAACCTGTGTCGACGCTGCATTAATGAAT	2848		
QY	2901	CGGCCAAACCGCGGGAGAGCGGTTTCGCTATTTGGCGCTCTTCGCTTCCGCTCAC	2960		
Db	2847	CGGCCAAACCGCGGGAGAGCGGTTTCGCTATTTGGCGCTCTTCGCTTCCGCTCAC	2788		
QY	2961	TGACTCGCTCGCTCGCTGCTTCGCGTGGCGAGCGGTATCAGCTCACTCAAAGCGGT	3020		
Db	2787	TGACTCGCTCGCTCGCTGCTTCGCGTGGCGAGCGGTATCAGCTCACTCAAAGCGGT	2728		
QY	3021	AATACGGTTATCACAGATCAGGGATTAACGAGGAAAGACATGTAGCAAAAGGCCA	3080		
Db	2727	AATACGGTTATCACAGATCAGGGATTAACGAGGAAAGACATGTAGCAAAAGGCCA	2668		
QY	3081	GCAAAAGGCCAGGAACCGTAAAAAGCCGGTGTGCTGGCGTTTTTCCATAGGCTCGGCC	3140		
Db	2667	GCAAAAGGCCAGGAACCGTAAAAAGCCGGTGTGCTGGCGTTTTTCCATAGGCTCGGCC	2608		
QY	3141	CCCTGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGGAAACCCGACAGACT	3200		
Db	2607	CCCTGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGGAAACCCGACAGACT	2548		
QY	3201	ATAAGATACAGCGCTTTCGCCCTGGAAGCTCCCTGTCGCTCTCTGTTCCGACCT	3260		
Db	2547	ATAAGATACAGCGCTTTCGCCCTGGAAGCTCCCTGTCGCTCTCTGTTCCGACCT	2488		
QY	3261	GCCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTGGCGTTTCTCAATG	3320		
Db	2487	GCCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTGGCGTTTCTCAATG	2428		
QY	3321	CTCAGCTGTAGGTATCTCAGTTCCGTTAGGTGCTGCTCCAGCTGGGCTGTGCA	3380		
Db	2427	CTCAGCTGTAGGTATCTCAGTTCCGTTAGGTGCTGCTCCAGCTGGGCTGTGCA	2368		
QY	3381	CGAACCCCGCTTCCAGCCCGACCGCTCGGCTTATCCGGTAACTATCGTTCGAGTCCAA	3440		
Db	2367	CGAACCCCGCTTCCAGCCCGACCGCTCGGCTTATCCGGTAACTATCGTTCGAGTCCAA	2308		
QY	3441	CCCGGTGAAGACACGACTTATCGCACTGGCAGCGACCTGGTAAACGAGTATAGCAGC	3500		
Db	2307	CCCGGTGAAGACACGACTTATCGCACTGGCAGCGACCTGGTAAACGAGTATAGCAGC	2248		
QY	3501	GAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTTAACTACCGCTACACTAG	3560		
Db	2247	GAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTTAACTACCGCTACACTAG	2188		
QY	3561	AAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGATTGG	3620		
Db	2187	AAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGATTGG	2128		
QY	3621	TAGCTCTTGATCCGGCAACAAACACCGCTGGTGGTGGTTTTTGTGTTGCAAGCA	3680		

Db	2127	TAGCTCTTGATCCGGCAACAAACACCGCTGGTGGTGGTTTTTGTGTTGCAAGCA	2068
QY	3681	GCAGATTACCGCGCAGAAAAAAGGATCTCAAGAGATCTTTTGATCTTTTCTACGGGTC	3740
Db	2067	GCAGATTACCGCGCAGAAAAAAGGATCTCAAGAGATCTTTTGATCTTTTCTACGGGTC	2008
QY	3741	TGACGCTCAGTGGAGCGAAACCTCAGTTAAGGATTTTGGTCATGAGTATTCGTG--	3798
Db	2007	TGACGCTCAGTGGAGCGAAACCTCAGTTAAGGATTTTGGTCATGAGTATTCGTGCTCA	1948
QY	3799	-----	3798
Db	1947	TAACTCTGCTTGGGGTCATTTATAGCGATTTTTCGGTATATCCATCTTTTTCGCACGA	1888
QY	3799	-----	3798
Db	1887	TATACAGGATTTTGC AAAAGGTTTCGTGTAGACTTTTCTTTGGTGTATCCAAACGCGGTCA	1828
QY	3799	-----ACCAAAGCGCCATCGTCCCTCCCACTCTGCAAGTTTCGGGGCATGGATGCGC	3852
Db	1827	CGGGCAGGATAGGTGAAGTAGGCCCAACCCGCGAGCGGGTGTCTTCTTCACTGTCCCT	1768
QY	3853	GGATAGCGCTGTCTGTGTTTCTCTGGATGCCGACGGAATTTGCATGCGCGGTAGAACTCCGC	3911
Db	1767	TATTCGACCTGGCGGTCTCAACGGGAATCTCTCTCGAGGCTGGCGGCTACCGCC	1708
QY	3912	-----GAGTCTGTCAGCTCAGGCTCAGGCTGAAACCACTCCGAGGGGATC	3959
Db	1707	GGCGTAAACAGATGAGGCAAGCGGATGGTGAATGAACCAAGCAACAGAGAGGCGAGC	1648
QY	3960	GAGCCCGGGTGGCGGAAGAACTCCAGCATGAGATCCCGCTGGAGGATCATCCAGC	4018
Db	1647	CCACTATCAAGGTGTACTGCTTCCAGACGAACGAAGAGCGATTGAGGAAAGCGCGC	1588
QY	4019	-----	4018
Db	1587	CGCGCCGCGCATGAGCTGTTCGGCTTACCTGTGCTGCGCTCAGGCGCTACAAATCAGC	1528
QY	4019	-----CGCGCTCCCGGAAACAGATTCGGAAGCCCAA	4049
Db	1527	GGCGTCTGTGACTATGAGCAGCTCCGCGAGGGCGTCCCGGAAACGATTCGGAAGCCCAA	1468
QY	4050	CCTTTCATAGAAGCGCGGTGGAATCGAAATCTCTGTGATGGCAGGTTGGCGTCGCTTG	4109
Db	1467	CCTTTCATAGAAGCGCGGTGGAATCGAAATCTCTGTGATGGCAGGTTGGCGTCGCTTG	1408
QY	4110	GTGCTCATTTTCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAAGCGGATAGAAG	4169
Db	1407	GTGCTCATTTTCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAAGCGGATAGAAG	1348
QY	4170	GCGATGCGCTCGGAATCGGGAGCGGCGATACCGTAAAGCAGAGGAGCGGTACGCCAT	4229
Db	1347	GCGATGCGCTCGGAATCGGGAGCGGCGATACCGTAAAGCAGAGGAGCGGTACGCCAT	1288
QY	4230	TCGCGCCCAAGCTCTTCAGCAATATCAGGTCAGCAGCTATGTCTCTGATAGCGGTCC	4289
Db	1287	TCGCGCCCAAGCTCTTCAGCAATATCAGGTCAGCAGCTATGTCTCTGATAGCGGTCC	1228
QY	4290	GCCACACCCAGCCGCGCCACAGTTCGATGAATCCAGAAAGCGGCCATTTTCCACCATGATA	4349
Db	1227	GCCACACCCAGCCGCGCCACAGTTCGATGAATCCAGAAAGCGGCCATTTTCCACCATGATA	1168
QY	4350	TTCCGCAAGCGGATCCCATATGGGTTCAGCAGAGATCTTCGCGTGGGATCGGGCC	4409
Db	1167	TTCCGCAAGCGGATCCCATATGGGTTCAGCAGAGATCTTCGCGTGGGATCGGGCC	1108
QY	4410	TTGAGCCTGGGAAACAGTTTCGGCTGGCGGAGCCCTGATGCTCTTCTCCAGATCATCC	4469
Db	1107	TTGAGCCTGGGAAACAGTTTCGGCTGGCGGAGCCCTGATGCTCTTCTCCAGATCATCC	1048
QY	4470	TGATCGAACAAGCCGGCTTCCATCCAGTACGTCGCTCGATGCGGATGTTTTCGTTGG	4529
Db	1047	TGATCGAACAAGCCGGCTTCCATCCAGTACGTCGCTCGATGCGGATGTTTTCGTTGG	988

4530 TGGTCGAATGGGAGGAGTACCGGATCAAGCGTATGACGCGCGCGCATTTGATCAGCATG 4589
Db TGGTCGAATGGGAGGAGTACCGGATCAAGCGTATGACGCGCGCGCATTTGATCAGCATG 928
QY ATGATATCTTTCTCGGAGGAGCAAGGTGATGATGACAGGAGATCTTCCCGCGGCACTTCG 4649
Db ATGATATCTTTCTCGGAGGAGCAAGGTGATGATGACAGGAGATCTTCCCGCGGCACTTCG 868
QY CCCAATAGCAGCCAGTCCCTTCCGCTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 4709
Db CCCAATAGCAGCCAGTCCCTTCCGCTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 808
QY ACGCCGTCGTCGCGCAGCAGCATGACGCGCTCCCTCGCTCGCTCGCTCGCTCGCTCGCTCG 4769
Db ACGCCGTCGTCGCGCAGCAGCATGACGCGCTCCCTCGCTCGCTCGCTCGCTCGCTCGCTCG 748
QY CCGGACAGTCCGCTTTCGACAAAAGAACCGGCGCGCGCTCGCTCGCTCGCTCGCTCGCTCG 4829
Db CCGGACAGTCCGCTTTCGACAAAAGAACCGGCGCGCGCTCGCTCGCTCGCTCGCTCGCTCG 688
QY CGGCGATCAGAGCAGCCGATGCTGTTGTCGCCAGTCATAGCCGCAATAGCTCTCCACC 4889
Db CGGCGATCAGAGCAGCCGATGCTGTTGTCGCCAGTCATAGCCGCAATAGCTCTCCACC 628
QY CGGCGATCAGAGCAGCCGATGCTGTTGTCGCCAGTCATAGCCGCAATAGCTCTCCACC 4949
Db CGGCGATCAGAGCAGCCGATGCTGTTGTCGCCAGTCATAGCCGCAATAGCTCTCCACC 568
QY CCGTGTCTTTGATCAGATCTTGCCTGCAATCCATCTTGTTCATCATGCGAAACGATCCTCAT 4950
Db CCGTGTCTTTGATCAGATCTTGCCTGCAATCCATCTTGTTCATCATGCGAAACGATCCTCAT 508
QY ATCCAGTTTACTTTTCAGCGGCTTCCCAACCTTACCAGAGGCGCGCCAGCTGGCAATTC 5069
Db ATCCAGTTTACTTTTCAGCGGCTTCCCAACCTTACCAGAGGCGCGCCAGCTGGCAATTC 448
QY GTTTCGCTGCTGTCATAAACCGGCGCATAGCTATGCGCATGCGCATGCGCATGCGCATGCG 5129
Db GTTTCGCTGCTGTCATAAACCGGCGCATAGCTATGCGCATGCGCATGCGCATGCGCATGCG 388
QY GTTTCGCTGCTGTCATAAACCGGCGCATAGCTATGCGCATGCGCATGCGCATGCGCATGCG 5189
Db GTTTCGCTGCTGTCATAAACCGGCGCATAGCTATGCGCATGCGCATGCGCATGCGCATGCG 328
QY TTATCCGGGGTCAGACCGTTTTCGCGACTGGGCTTTCTAGCTTTCCGCTTCCCTTTAG 5249
Db TTATCCGGGGTCAGACCGTTTTCGCGACTGGGCTTTCTAGCTTTCCGCTTCCCTTTAG 268
QY CAGCCCTTGGCCCTGAGTGTGTTGGGCGAGCGTG 5283
Db CAGCCCTTGGCCCTGAGTGTGTTGGGCGAGCGTG 234

RESULT 15
AY222811/c 3053 bp DNA circular SYN 24-MAR-2003
LOCUS Cloning vector pK18PolyF2, complete sequence.
DEFINITION
ACCESSION AY222811
VERSION AY222811.1 GI:29169334
KEYWORDS
SOURCE
ORGANISM
Cloning vector pK18PolyF2
Cloning vector pK18PolyF2
artificial sequences; vectors.
REFERENCE
1 (bases 1 to 3053)
AUTHORS Tauch,A., Kirchner,O., Löffler,B., Gotker,S., Puhler,A. and Kalinowski,J.
TITLE Efficient electrotransformation of Corynebacterium diphtheriae with a mini-replicon derived from the Corynebacterium glutamicum plasmid pGAI
JOURNAL Curr. Microbiol. 45 (5), 362-367 (2002)
MEDLINE 22217619
PUBMED 12232668

2 (bases 1 to 3053)
Tauch,A., Kirchner,O., Löffler,B., Gotker,S., Puhler,A. and Kalinowski,J.
Direct Submission
Submitted (22-JAN-2003) Department of Genetics, University of Bielefeld, Universitätsstrasse 25, Bielefeld D-33615, Germany
Location/Qualifiers
FEATURES
source
1..3053
/organism="Cloning vector pK18PolyF2"
/mol_type="genomic DNA"
/db_xref="taxon:224892"
/lab_host="Escherichia coli"
602..1396
/gene="aph(3')-IIa"
602..1396
/gene="aph(3')-IIa"
/note="aminoglycoside phosphotransferase"
/codon_start=1
/transl_table=1
/product="Aph(3')-IIa"
/protein_id="AAO65976.1"
/db_xref="GI:29169335"
/translation="MIEQDGLHAGSPAAWVERLFYDWAQQTIGSDAAVRLSAQGR
PVLFKYTLGALNELODEAARLWLTAVLPVCAVLDVTVTEAGRDWLLGVEPQDL
LSSHLAPAEKYSIMADAMRLHLDPATCFPDHOKHRIERATRMEAGLVQDDLE
BHQGLAPRELPAELKARMPDGEDLVVTHGDACLNIMVENGFRSGFIDCRLGVADRY
BQIALATRDIAELGGEWADFLVLYGIAAPDSQRIAFYRLLDSEF"
misc_feature
1781..1879
/note="RP4-mob region"
2642..2644
/note="oriV"
BASE COUNT 673 a 842 c 874 g 664 t
ORIGIN
Query Match 39.3%; Score 2078; DB 12; Length 3053;
Best Local Similarity 86.7%; Pred. No. 0;
Matches 2423; Conservative 0; Mismatches 140; Indels 231; Gaps 3;
QY 2721 GTCATAGCTGTTTCTGTCGTGAAATTTGTTATCCGCTCACAATTCACACAAATAGGAGC 2780
Db 3045 GTCATAGCTGTTTCTGTCGTGAAATTTGTTATCCGCTCACAATTCACACAAATAGGAGC 2986
QY 2781 CGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTTAATGAGTGAGTAACTACATTAATTCG 2840
Db 2985 CGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTTAATGAGTGAGTAACTACATTAATTCG 2926
QY 2841 GTTCCGCTCACTGCGCGCTTCCAGTCGCGAAGACCTGTCGTCAGCTGCATTAAATGAAT 2900
Db 2925 GTTCCGCTCACTGCGCGCTTCCAGTCGCGAAGACCTGTCGTCAGCTGCATTAAATGAAT 2866
QY 2901 CGGCCAAACGCGCGGAGAGCGGTTTGGCTATTGGCGGCTTTCCGCTTCTCGCTCAC 2960
Db 2865 CGGCCAAACGCGCGGAGAGCGGTTTGGCTATTGGCGGCTTTCCGCTTCTCGCTCAC 2806
QY 2961 TGACTCGCTCGCTCGGTCGTTCCGCTGGCGAGCGGTATCAGTCACTCAAGAGCGGT 3020
Db 2805 TGACTCGCTCGCTCGGTCGTTCCGCTGGCGAGCGGTATCAGTCACTCAAGAGCGGT 2746
QY 3021 AATACGGTTATCCACAGAATCAGGGGATAACCGAGGAAAGAACATGTGAGCAAAAGGCCA 3080
Db 2745 AATACGGTTATCCACAGAATCAGGGGATAACCGAGGAAAGAACATGTGAGCAAAAGGCCA 2686
QY 3081 GCAAAAGCCAGGAACCGTAAAAAGCGCGGTTTGGCGGTTTTCATAGGCTCCGCC 3140
Db 2685 GCAAAAGCCAGGAACCGTAAAAAGCGCGGTTTGGCGGTTTTCATAGGCTCCGCC 2626
QY 3141 CCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGACT 3200
Db 2625 CCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGACT 2566
QY 3201 ATAAAGATACAGCGGTTTCCCGCTGGAGCTCCCTCGCTCGCTTCTCTTCCGACCT 3260
Db 2565 ATAAAGATACAGCGGTTTCCCGCTGGAGCTCCCTCGCTCGCTTCTCTTCCGACCT 2506

QY	3261	GGCGCTTACGGATACCTGTGCGCCCTTTCTCCCTTCGGAAAGCTGGCGCTTTCTCATAG	3320
Db	2505	GCCGCTTACGGATACCTGTGCGCCCTTTCTCCCTTCGGAAAGCTGGCGCTTTCTCAATG	2446
QY	3321	CTCACGCTGTAGGTATCTCAGTTCGGTGTAGTTCGTTCCCTCCAAAGCTGGCTGTGTGCA	3380
Db	2445	CTCACGCTGTAGGTATCTCAGTTCGGTGTAGTTCGTTCCCTCCAAAGCTGGCTGTGTGCA	2386
QY	3381	CGAAACCCCGCTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTTGAGTCCAA	3440
Db	2385	CGAAACCCCGCTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTTGAGTCCAA	2326
QY	3441	CCCGGTTAAGCACGACTTATCGCCACTGCGACGAGCAGCACTGGTTAACAGATTAGCAGAGC	3500
Db	2325	CCCGGTTAAGCACGACTTATCGCCACTGCGACGAGCAGCACTGGTTAACAGATTAGCAGAGC	2266
QY	3501	GAGGTATGTAGGCGGTGCTACAGAGTCTTTGAAGTGTGGCCTAACTACGGCTACACTAG	3560
Db	2265	GAGGTATGTAGGCGGTGCTACAGAGTCTTTGAAGTGTGGCCTAACTACGGCTACACTAG	2206
QY	3561	AAGAAACAGTATTTGGTATCTGCGCTCTGTGAAGCCAGATTACTCTCGGAAGAGAGTTCG	3620
Db	2205	AAGAAACAGTATTTGGTATCTGCGCTCTGTGAAGCCAGATTACTCTCGGAAGAGAGTTCG	2146
QY	3621	TAGCTCTTGATCCGGCAAAACAAACCCGCTGCTAGCGGTGGTTTTTTGTTTGCAGCA	3680
Db	2145	TAGCTCTTGATCCGGCAAAACAAACCCGCTGCTAGCGGTGGTTTTTTGTTTGCAGCA	2086
QY	3681	GCAGATTACCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGCTC	3740
Db	2085	GCAGATTACCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGCTC	2026
QY	3741	TGACGCTCAGTGAAACGAAACTCAGGTTAAGGGATTTTGGTCAATGAGATTATCGTCG--	3798
Db	2025	TGACGCTCAGTGAAACGAAACTCAGGTTAAGGGATTTTGGTCAATGAGATTATCGTCG--	1966
QY	3799	-----	3798
Db	1965	TAAACCTGTCTCGGGTCAATTATAGCGATTTTTCGGTATATCCATCCCTTTTTCGCACGA	1906
QY	3799	-----	3798
Db	1905	TATACAGGATTTTGCACAAAGGGTTCGTGTAGACTTTCTTTGGTGTATCCAAACGGCGTCAG	1846
QY	3799	-----ACAAAGGGCCATCTGTGCTCCCACTCTGTGAGTTTCGGGGGATCGATGCGC	3852
Db	1845	CGGGGACAGGATAGTGAAGTAGGAGGCCCAACCCGCGAGCGGGTGTTCCTTTCTCATGTGCCCT	1786
QY	3853	GGATAGCGCTGTGTGTTCTCTGGATCCGCACGGATTTGCACTGCCGGTAGAAACTCCGC-	3911
Db	1785	TATTCGACCTGCGGTGCTCAACGGGAATCTGCTCTGCGAGSCTGGCGCGCTACCGCC	1726
QY	3912	-----GAGGTGCTCAGCCTCAGCAGCAGCTGAAACCACTCCGAGGGGATC	3959
Db	1725	GGCGTAAACAGATGAGGGCAACGCGATGGCTGATGAAACCAAGCCCAACAGGAAGGGCAGC	1666
QY	3960	GAGCCCGGGGTGGCGAAGAACTCCAGCATGAGATCCCGCGCTGGAGGATCATCCAGC-	4018
Db	1665	CCACCTATCAAGGTGTACTGCTCTCCAGACGAACGAGAGCGGATTGAGGAAAGGCGCGC	1606
QY	4019	-----	4018
Db	1605	CGCGCCCGCATGAGCCTGTCCGCTACCTGCTGCGCGCTCGGCCAGGCTACAAAATCACG	1546
QY	4019	-----CGCGCTCCCGGAAAAACGATTCGGAAGCCCAA	4049
Db	1545	GGCGTGTGACTATGACGACGTCGCGAGGGGCTCCCGGAACAGATTCCGAAGCCCAA	1486
QY	4050	CTTTTCATAGAAGCGCGGTGGAAATCGAAATCTCGTGTATGGCAGGTTGGCGTCGCTTG	4109
Db	1485	CTTTTCATAGAAGCGCGGTGGAAATCTCGTGTATGGCAGGTTGGCGTCGCTTG	1426

QY	4110	GTCCGTCATTTCGAACCC	CAGAGTCCCGCTCAGAGAA	CTCGTCAAGAAAGCGGATAGAAG	4166
DB	1425	GTCCGTCATTTCGAACCC	CAGAGTCCCGCTCAGAGAA	CTCGTCAAGAAAGCGGATAGAAG	1366
QY	4170	GGGATGCGTGGGAATCGG	AGCGGATACCGTAAAGCAC	GAGGAAGCGGTTCAGCCCAT	4229
DB	1365	GGATGCGTGGGAATCGG	AGCGGATACCGTAAAGCAC	GAGGAAGCGGTTCAGCCCAT	1306
QY	4230	TGCGCGCCAAAGCTCTTT	CAGCAATATCACGGGTAG	CAACGCTATGCTTGATAGCGGTCC	4289
DB	1305	TGCGCGCCAAAGCTCTTT	CAGCAATATCACGGGTAG	CAACGCTATGCTTGATAGCGGTCC	1246
QY	4290	GGCACACCCAGCGGCGCA	CAGTCGATGAATCCAGAAA	AGCGGCCATTTTTCACCACTGATA	4349
DB	1245	GGCACACCCAGCGGCGCA	CAGTCGATGAATCCAGAAA	AGCGGCCATTTTTCACCACTGATA	1186
QY	4350	TTGCGCAAGCAGCGCATPC	GGCATGGGTACGACGAGAT	CTCGCGCTCGGGCATTCGCGCGC	4409
DB	1185	TTGCGCAAGCAGCGCATPC	GGCATGGGTACGACGAGAT	CTCGCGCTCGGGCATTCGCGCGC	1126
QY	4410	TTGAGCCTGGCGNAACAG	TTTCGGCTGCGGAGCGCC	CTGATGCTCTTCGTCAGATCATCC	4469
DB	1125	TTGAGCCTGGCGNAACAG	TTTCGGCTGCGGAGCGCC	CTGATGCTCTTCGTCAGATCATCC	1066
QY	4470	TGATCCAGCAAGACCGG	CTTCCATCCGAGTACGTG	CTCGCTCGATGCGATGCTTCGTTGG	4529
DB	1065	TGATCCAGCAAGACCGG	CTTCCATCCGAGTACGTG	CTCGCTCGATGCGATGCTTCGTTGG	1006
QY	4530	TGCTCGAATGGGGAGGT	AGCCGATCAAGGGTATG	CAGCGCGCGCATTCGATCAGCATG	4589
DB	1005	TGCTCGAATGGGGAGGT	AGCCGATCAAGGGTATG	CAGCGCGCGCATTCGATCAGCATG	946
QY	4590	ATGGATACTTTTCGCGA	GAGCAAGGTGAGATGAC	GAGAGATCTTCGCCCGGCATCTCG	4649
DB	945	ATGGATACTTTTCGCGA	GAGCAAGGTGAGATGAC	GAGAGATCTTCGCCCGGCATCTCG	886
QY	4650	CCCAATAGCAGCGAGTCC	CTTCCGCTTCAGTGACAA	CGTTCGAGCACAGTTCGCGCAAGGA	4709
DB	885	CCCAATAGCAGCGAGTCC	CTTCCGCTTCAGTGACAA	CGTTCGAGCACAGTTCGCGCAAGGA	826
QY	4710	ACGCCGCTGTGCCAGC	ACGATAGCCGCTGCTCT	CTGTCAGTTTCATTTCAGGGCA	4769
DB	825	ACGCCGCTGTGCCAGC	ACGATAGCCGCTGCTCT	CTGTCAGTTTCATTTCAGGGCA	766
QY	4770	CCGGACAGGTTCGTCTTG	ACAAAAGAACCGGCGC	CGCCCTCGCTGACAGCCGGAAACAG	4829
DB	765	CCGGACAGGTTCGTCTTG	ACAAAAGAACCGGCGC	CGCCCTCGCTGACAGCCGGAAACAG	706
QY	4830	GGCGCATCAGACAGCGG	ATGCTGCTTGCCCATG	CTATAGCCGAATAGCCTCTCCACC	4889
DB	705	GGCGCATCAGACAGCGG	ATGCTGCTTGCCCATG	CTATAGCCGAATAGCCTCTCCACC	646
QY	4890	CAAGCGGCCGGGAAAC	CTGCGTGCATTCATCT	TGTTCAATCATGCGAAACGATCTCTCAT	4949
DB	645	CAAGCGGCCGGGAAAC	CTGCGTGCATTCATCT	TGTTCAATCATGCGAAACGATCTCTCAT	586
QY	4950	CTGTCTCTTGATCAGAT	CTTGATCCGCTGCGCAT	CAGATCCCTTGGCGCGAGGAAGCC	5009
DB	585	CTGTCTCTTGATCAGAT	CTTGATCCGCTGCGCAT	CAGATCCCTTGGCGCGAGGAAGCC	526
QY	5010	ATCCAGTTTACTTTGAG	GGGCTTTCCAAACCTT	ACAGAGGGCGCCCGCATGTGGCAATTC	5069
DB	525	ATCCAGTTTACTTTGAG	GGGCTTTCCAAACCTT	ACAGAGGGCGCCCGCATGTGGCAATTC	466
QY	5070	GGTTCGCTGCTGCAT	TAATAAACCGGCCAG	TCTAGCTATGCGCATGTAAGCCCACTGCAA	5129
DB	465	GGTTCGCTGCTGCAT	TAATAAACCGGCCAG	TCTAGCTATGCGCATGTAAGCCCACTGCAA	406
QY	5130	GCTACCTGCTTCTCT	TTTGCGCTTGCGTTT	CTCTTGCTCCAGATAGCCCACTAGCTGACA	5189
DB	405	GCTACCTGCTTCTCT	TTTGCGCTTGCGTTT	CTCTTGCTCCAGATAGCCCACTAGCTGACA	346
QY	5190	TTCATCCGCGGTCAG	ACCGTTTCTCGGAG	CTGGCTTTCACGTTTCGCGTTCCTTTAG	5249

Db 345 TTATCCGGGTCAGACCGTTTCTGCGACTGGCTTTCTACGTGTTCCGCTTCCTTAG 286

QY 5250 CAGCCCTTGGCCCTGAGTGTCTTGGCGCAGCGTG 5283

Db 285 CAGCCCTTGGCCCTGAGTGTCTTGGCGCAGCGTG 252

Search completed: February 15, 2004, 16:21:29
Job time : 18649 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 05:44:19 ; Search time 1245 Seconds
(without alignments)

11454.724 Million cell updates/sec

Title: US-09-921-143-36

Perfect score: 5283

Sequence: 1 aagttgacctatgcgact.....tgagtgttcggcgcggtg 5283

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 19Jun03.*

```

1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5283	100.0	5283	24	ABK10062 Expression vector
2	2240	42.4	8349	21	AA15565 pMON30464 plasmid.
3	2017.2	38.2	6136	19	AAV29673 Hybrid NAMK.6 (ter
C 4	1793.4	33.9	4800	20	AAV52020 Synthetic DNA plas
C 5	1652.4	31.3	5594	19	AAV33629 GENSA 981, a monom
C 6	1634.4	30.9	5285	22	AAH74865 Nucleotide sequenc
C 7	1618.2	30.6	6561	19	AAV33630 Plasmid VLSNO2 DNA
8	1603.8	30.4	2192	24	ABQ78294 Nucleotide sequenc

9	1603.8	30.4	3534	24	ABQ78293	Nucleotide sequenc
10	1603.8	30.4	3599	19	AAV50428	Plasmid pIG0552 lo
11	1603.8	30.4	3599	19	AAV40796	Actual sequence of
12	1602.2	30.3	3600	19	AAV50427	Plasmid pIG0552 up
13	1602.2	30.3	3600	19	AAV40795	Expected sequence
14	1602.2	30.3	3600	19	AAV40795	Plasmid pIG0335 DN
15	1600	30.3	3566	20	AAZ40417	Plasmid pIN0961 co
C 16	1571.6	29.7	4665	22	AAF30342	Construct pBIOPE-
C 17	1570.8	29.7	5068	20	AAH84027	MP9 promoter GFP
18	1570	29.7	4517	22	AAH24427	pBKMV phagemid.
19	1570	29.7	4518	19	AAV33626	Plasmid pBK-CMV DN
20	1570	29.7	4886	24	AAH24427	Human aequorin-pBk
C 21	1570	29.7	5031	20	AAZ23355	EGFP/DRM fusion pr
C 22	1570	29.7	5034	20	AAZ23354	EGFP/DRM fusion pr
C 23	1570	29.7	5130	20	AAZ23353	EGFP/DRM fusion pr
C 24	1570	29.7	5168	20	AAZ23351	EGFP/DRM fusion pr
C 25	1570	29.7	5619	21	AAA07773	DNA sequence of pl
C 26	1570	29.7	5627	21	AAA07778	DNA sequence of pl
C 27	1570	29.7	5775	25	ABZ24716	Murine potassium c
C 28	1570	29.7	6121	21	AAA07779	DNA sequence of pl
C 29	1570	29.7	6139	21	AAZ40023	Interleukin-12 fus
C 30	1570	29.7	7655	21	AAA07847	Hs-UNC-53/1 fragm
C 31	1570	29.7	8140	21	AAA07845	Hs-UNC-53/3 fragm
C 32	1570	29.7	11842	21	AAA07846	Hs-UNC-53/3 DNA c
C 33	1569.8	29.7	4055	22	AAI68460	Plasmid pSV-HIII-C
C 34	1569.8	29.7	4853	22	AAI10015	Plasmid pSFM 8 to
C 35	1569.8	29.7	4892	24	AAI47954	Modified yeast str
C 36	1569.8	29.7	4892	25	ACC44640	Vector pDRedIN1 n
C 37	1569.8	29.7	4895	21	AAA50945	Anemonia majano mu
C 38	1569.8	29.7	4700	21	AAA48561	Zoanthus sp. novel
C 39	1569.8	29.7	4700	21	AAA48562	Zoanthus sp. novel
C 40	1569.8	29.7	4727	25	ABQ80035	VSV M C-terminal/H
C 41	1569.8	29.7	4729	21	AAA07772	DNA sequence of ve
C 42	1569.8	29.7	4735	21	AAA07777	DNA sequence of ve
C 43	1569.8	29.7	4740	25	ABQ80036	Hifi-alpha ODD/VS
C 44	1569.8	29.7	4821	21	AAA50946	Anemonia majano mu
C 45	1569.8	29.7	4821	21	AAA50947	Anemonia majano mu

ALIGNMENTS

RESULT 1
ABK10062
ID ABK10062 standard; DNA; 5283 BP.
XX
XX ABK10062;
AC
XX
XX 21-MAY-2002 (first entry)
DT
XX
XX Expression vector construct pVGI.1 containing VEGF-2 insert.
DE
XX
XX Human; vascular endothelial growth factor 2; VEGF-2; ds; pVGI.1;
KW chronic limb ischaemia; myocardial ischaemia; autoimmune disorder; pHE4a;
KW allergic reaction; organ rejection; inflammatory condition; arrhythmia;
KW hyperproliferative disorder; viral infection; bacterial infection;
KW fungal infection; parasitic infection; cardiovascular disorder; embolism;
KW heart valve disease; aneurysm; arterial occlusive disorder; gene therapy.
OS
XX Synthetic.
XX
XX WO200211769-A1.
XX
XX 14-FEB-2002.
PD
XX
XX 03-AUG-2001; 2001WO-US24658.
PF
XX
XX 04-AUG-2000; 2000US-223276P.
PR
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Coleman TA;
PI
XX

DR WPI; 2002-217153/27.
XX Isolated nucleic acid having expression vector construct with vascular
PT endothelial growth factor-2 insert, useful for treating chronic limb
PT ischaemia or myocardial ischaemia, autoimmune disorders and allergic
PT conditions -
XX
PS Disclosure; Fig 31; 241pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising pVGI.1
CC expression vector construct containing the vascular endothelial growth
CC factor 2 (VEGF-2) insert. The nucleic acid is useful for producing a host
CC cell by transducing, transforming or transfecting a host cell with the
CC DNA and for treating a patient having chronic limb ischaemia or
CC myocardial ischaemia, or a disease or disorder selected from autoimmune
CC disorders (e.g. rheumatoid arthritis, dermatitis), allergic reactions or
CC conditions (e.g. asthma), organ rejection, inflammatory conditions (e.g.
CC Crohn's disease), hyperproliferative disorders (e.g. Gaucher's disease),
CC diseases due to viral, bacterial, fungal or parasitic infection,
CC cardiovascular disorders (e.g. cardiomyopathy), arrhythmia, heart valve
CC diseases, aneurysms, arterial occlusive disorders and embolism. This
CC sequence represents the pVGI.1 expression vector containing the VEGF-2
CC insert.
XX
SQ Sequence 5283 BP; 1244 A; 1422 C; 1417 G; 1200 T; 0 other;

Query Match 100.0%; Score 5283; DB 24; Length 5283;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGCTTGACCTTATGGACCTTCTCTACTTGGCAGTACATCTAGCTATTAGTATCATCGCTAT 60
DB 1 AAGCTTGACCTTATGGACCTTCTCTACTTGGCAGTACATCTAGCTATTAGTATCATCGCTAT 60

QY 61 TACCATGGTGATCGGTTTGGCAGTACATCAATGGGCGTGGATACGGTTTGACTCAG 120
DB 61 TACCATGGTGATCGGTTTGGCAGTACATCAATGGGCGTGGATACGGTTTGACTCAG 120

QY 121 GGGATTTCCAAAGTCTCCACCCCTACGCTGATGGAGTTTGTGTTGGCACCACCAATCA 180
DB 121 GGGATTTCCAAAGTCTCCACCCCTACGCTGATGGAGTTTGTGTTGGCACCACCAATCA 180

QY 181 ACGAGACTTTCCAAATGTCTGAACAACTCCGCCCATTTGACGCAATGGGCGGTAGGCG 240
DB 181 ACGAGACTTTCCAAATGTCTGAACAACTCCGCCCATTTGACGCAATGGGCGGTAGGCG 240

QY 241 AACATGCTTATGTAAACGGTGAGTTAGCAATGCTTATAGGAGAGAGAAAGCACCGTG 300
DB 241 AACATGCTTATGTAAACGGTGAGTTAGCAATGCTTATAGGAGAGAGAAAGCACCGTG 300

QY 301 CATGCCGATTTGGTGGGAGTAAAGTGGTATGATCGTGTATGATCGTGCCTTTGTTAGGAAG 360
DB 301 CATGCCGATTTGGTGGGAGTAAAGTGGTATGATCGTGTATGATCGTGCCTTTGTTAGGAAG 360

QY 361 GCAACAGACGGGTCTAACCGATTTGACGACCACTGATTCGCAATTCGCAATTCGAGATATT 420
DB 361 GCAACAGACGGGTCTAACCGATTTGACGACCACTGATTCGCAATTCGCAATTCGAGATATT 420

QY 421 GTATTTAAGTCCCGAGCTCCGATACAAATAAGCGCCATTGACCAATTGACCAATTGGTGTG 480
DB 421 GTATTTAAGTCCCGAGCTCCGATACAAATAAGCGCCATTGACCAATTGACCAATTGGTGTG 480

QY 481 CACTCGGTTGGGATTCGATTCGATCATGCACTGCTGGGCTTCTCTCTGCGGCTGTCT 540
DB 481 CACTCGGTTGGGATTCGATTCGATCATGCACTGCTGGGCTTCTCTCTGCGGCTGTCT 540

QY 541 CTGCTCGCGCTGCGCTGCTCCCGGCTCTCGAGGCGCCCGCGCGCGCGCGCTTC 600
DB 541 CTGCTCGCGCTGCGCTGCTCCCGGCTCTCGAGGCGCCCGCGCGCGCGCGCTTC 600

QY 601 GAGTCCGACTCGACCTCTCGAGCGCGAGCGCCGAGCGCGGTAGGCGCACGGCTTAGCA 660
DB 601 GAGTCCGACTCGACCTCTCGAGCGCGAGCGCCGAGCGCGGTAGGCGCACGGCTTAGCA 660

QY 661 AGCAAGAATCTGAGGAGCAGTTACGGTCTGTGTCCAGTGTAGATCAATCATGATGTA 720
DB 661 AGCAAGAATCTGAGGAGCAGTTACGGTCTGTGTCCAGTGTAGATCAATCATGATGTA 720

QY 721 CTCTACCCAGATATTGGAAATGTACAAGTGCAGCTAAGGAAGAGGCTGGCAACAT 780
DB 721 CTCTACCCAGATATTGGAAATGTACAAGTGCAGCTAAGGAAGAGGCTGGCAACAT 780

QY 781 AACAGAGAACAGCCCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGTCGAGCA 840
DB 781 AACAGAGAACAGCCCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGTCGAGCA 840

QY 841 CATTATTAACAGAGATCTTGAAGATTTGATAAAGTGAATAGTGGAGAGAGTCAATGCATG 900
DB 841 CATTATTAACAGAGATCTTGAAGATTTGATAAAGTGAATAGTGGAGAGAGTCAATGCATG 900

QY 901 CCACGGAGGTGTGTATAGATGTGGGGAAGGAGCTTTGGAGTCCGACAAACACCTTTCTTT 960
DB 901 CCACGGAGGTGTGTATAGATGTGGGGAAGGAGCTTTGGAGTCCGACAAACACCTTTCTTT 960

QY 961 AAACCTCCATGTGTCTCCGCTACAGATGTGGGGGTGTGCTGCAATAGTGGGGGTGCGAG 1020
DB 961 AAACCTCCATGTGTCTCCGCTACAGATGTGGGGGTGTGCTGCAATAGTGGGGGTGCGAG 1020

QY 1021 TGCATGAACACACGACGACGTACCTCAGCAAGACGTTATTTCGAAATTAACAGTGCCTCTC 1080
DB 1021 TGCATGAACACACGACGACGTACCTCAGCAAGACGTTATTTCGAAATTAACAGTGCCTCTC 1080

QY 1081 TCTCAAGGCCCCAAACCCAGTAAACATCAGTTTTTGGCCATCACACTTCTCCGATGCGATG 1140
DB 1081 TCTCAAGGCCCCAAACCCAGTAAACATCAGTTTTTGGCCATCACACTTCTCCGATGCGATG 1140

QY 1141 TCTAAACTGGATGTTTACAGCAAGTTTCATTCATTTATTAGAGCTTCCTTCGCGAGCAACA 1200
DB 1141 TCTAAACTGGATGTTTACAGCAAGTTTCATTCATTTATTAGAGCTTCCTTCGCGAGCAACA 1200

QY 1201 CTACCAAGTGTCTAGGAGCGAAACAGCCTGCCCAACCAATTAATCATGTGGAAATATCATC 1260
DB 1201 CTACCAAGTGTCTAGGAGCGAAACAGCCTGCCCAACCAATTAATCATGTGGAAATATCATC 1260

QY 1261 ATCTGCGAGTCCGCTCAGGAGATTTTATGTTTTTCTCGGATGCTGGAGATGACTCA 1320
DB 1261 ATCTGCGAGTCCGCTCAGGAGATTTTATGTTTTTCTCGGATGCTGGAGATGACTCA 1320

QY 1321 ACAGATGGATTTCCATGACATCTGTGGACCAACCAAGGAGCTGGATGAAGAGACCTGTCTAG 1380
DB 1321 ACAGATGGATTTCCATGACATCTGTGGACCAACCAAGGAGCTGGATGAAGAGACCTGTCTAG 1380

QY 1381 TGTGTCTGAGAGCGGGCTTCGGCTTCGAGCTGTGGACCCCAAGAACTAGACAGA 1440
DB 1381 TGTGTCTGAGAGCGGGCTTCGGCTTCGAGCTGTGGACCCCAAGAACTAGACAGA 1440

QY 1441 AACTCATGCCAGTGTCTGTAAACCAAACTCTTCCCGAGCAATGTGGGGCCAAACCGA 1500
DB 1441 AACTCATGCCAGTGTCTGTAAACCAAACTCTTCCCGAGCAATGTGGGGCCAAACCGA 1500

QY 1501 GAATTTGATGAACACATGCCAGTGTGTATGAAGAGACCTTCGCCAGAAATCAACCC 1560
DB 1501 GAATTTGATGAACACATGCCAGTGTGTATGAAGAGACCTTCGCCAGAAATCAACCC 1560

QY 1561 CTAAATCTCGAAATGTGCTGTAATGTACAGAAAGTCCACAGAAATGCTTGTAAAAA 1620
DB 1561 CTAAATCTCGAAATGTGCTGTAATGTACAGAAAGTCCACAGAAATGCTTGTAAAAA 1620

QY 1621 GGAAGAAAGTCCACCAACCAACATGCGCTGTACAGAGGCGCATGTACGAACGGCCAG 1680
DB 1621 GGAAGAAAGTCCACCAACCAACATGCGCTGTGTACAGAGGCGCATGTACGAACGGCCAG 1680

QY 1681 AAGGCTCTGAGGACGAGGATTTTCATATAGTGAAGAGTGTCTGCTGTGTCTTCATAT 1740
DB 1681 AAGGCTCTGAGGACGAGGATTTTCATATAGTGAAGAGTGTCTGCTGTGTCTTCATAT 1740

1741 TGGAAAGACCAAAATGAGCTAATCTAGGATCGGTACCGTCCCTGCGCCAGCGCTTTTGTCAAAC 1800
1741 TGGAAAGACCAAAATGAGCTAATCTAGGATCGGTACCGTCCCTGCGCCAGCGCTTTTGTCAAAC 1800
1801 AGCACTTTTGTGGTCTCACTTGTGTGGAAGCTCTCTACCTGTGTGTGGAGAGGTGGAT 1860
1801 AGCACTTTTGTGGTCTCACTTGTGTGGAAGCTCTCTACCTGTGTGTGGAGAGGTGGAT 1860
1861 TCTTCTACACACCCATGTCCTGCGCGAAGTGGAGGACCCACAAGGTAAAGCTCTGCTCCT 1920
1861 TCTTCTACACACCCATGTCCTGCGCGAAGTGGAGGACCCACAAGGTAAAGCTCTGCTCCT 1920
1921 GAATTTCTATCCCAAGTGTAACTACCTCTGTTTGTCTTTCACCTTTGAGACCTTGTAAAT 1980
1921 GAATTTCTATCCCAAGTGTAACTACCTCTGTTTGTCTTTCACCTTTGAGACCTTGTAAAT 1980
1981 GTGCCCTAGGTGTGAGAGGTCTCAGGCTAACAGTGGGGGACATTTCTGTGGGAGCT 2040
1981 GTGCCCTAGGTGTGAGAGGTCTCAGGCTAACAGTGGGGGACATTTCTGTGGGAGCT 2040
2041 AGACATATGTAAACATGGTAGTGTCCAGGAAGGAGTGAAGATCTCTTAAAGTCTCCTA 2100
2041 AGACATATGTAAACATGGTAGTGTCCAGGAAGGAGTGAAGATCTCTTAAAGTCTCCTA 2100
2101 GGTGTGACGGGTGTGAGAGGTCTCAGGCTAACAGTGGGGGACATTTCTGTGGGAGCT 2160
2101 GGTGTGACGGGTGTGAGAGGTCTCAGGCTAACAGTGGGGGACATTTCTGTGGGAGCT 2160
2161 GCACCTGACTGAGGGATGTGAACAGGATGTGTAGTGTGGAGGGCCATATGTCATTCAT 2220
2161 GCACCTGACTGAGGGATGTGAACAGGATGTGTAGTGTGGAGGGCCATATGTCATTCAT 2220
2221 GACCAAGTACTTGTCTCAGCAGCATGMAACCTTGCCTCTCTGCTGACCTTACAGAGGGA 2280
2221 GACCAAGTACTTGTCTCAGCAGCATGMAACCTTGCCTCTCTGCTGACCTTACAGAGGGA 2280
2281 TAAAGTGAGAGAAAGCTGGGTGTATCAGGGGTGCTCAGCTCTCTCTTAACTGGATGT 2340
2281 TAAAGTGAGAGAAAGCTGGGTGTATCAGGGGTGCTCAGCTCTCTCTTAACTGGATGT 2340
2341 CCTATGTGCTTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
2341 CCTATGTGCTTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
2401 GCAGTGGCAAACTGGAGTGGGTGGAGCCCGGGGGCAGGTGACCTTCAGACCTTGGCA 2460
2401 GCAGTGGCAAACTGGAGTGGGTGGAGCCCGGGGGCAGGTGACCTTCAGACCTTGGCA 2460
2461 CTGGAGGTGGCCCGGCGAGAGCGCGCATCGTGGATCGATGCTGCAACGAGCATCTGCTCT 2520
2461 CTGGAGGTGGCCCGGCGAGAGCGCGCATCGTGGATCGATGCTGCAACGAGCATCTGCTCT 2520
2521 CTCTACAACTGGAGAACTACTGCACTAGGCGCCACCACTACCTGTCCACCCCTCTGCA 2580
2521 CTCTACAACTGGAGAACTACTGCACTAGGCGCCACCACTACCTGTCCACCCCTCTGCA 2580
2581 ATGAATAAACCCTTTCAGAGAGCACTACAAAGTGTGTGTACATGCGGTGATGTGCATATG 2640
2581 ATGAATAAACCCTTTCAGAGAGCACTACAAAGTGTGTGTACATGCGGTGATGTGCATATG 2640
2641 TGGTGGGGGGGAAACATGATGAGTGGGTGCTGAGTGGGTGCGGCTTAACTATCTGCGCA 2700
2641 TGGTGGGGGGGAAACATGATGAGTGGGTGCTGAGTGGGTGCGGCTTAACTATCTGCGCA 2700
2701 GCTGTCTAGAGCTAACTCATGCTCATAGCTGTTTCTGTGTGAAATTTGTTATCGGCTCACA 2760
2701 GCTGTCTAGAGCTAACTCATGCTCATAGCTGTTTCTGTGTGAAATTTGTTATCGGCTCACA 2760
2761 ATTCCACACAACTACAGCCGGAAGCATAAAGTGTAAAGCTTGGGGTGCCTTAATGAGTG 2820
2761 ATTCCACACAACTACAGCCGGAAGCATAAAGTGTAAAGCTTGGGGTGCCTTAATGAGTG 2820
2821 AGCTAACTCACAATAATTTGGTGTGGCTCACTGCGCGCTTCCAGTCCGGGAACCTGTGCG 2880

2821 AGCTAACTCACAATAATTTGGTGTGGCTCACTGCGCGCTTTCCAGTCCGGGAACCTGTGCG 2880
2881 TCCAGCTGCTAATTAATGAAATCGGCCAAAGCGCGGGAGAGCGGTGTTGCTATTGGCGCG 2940
2881 TCCAGCTGCTAATTAATGAAATCGGCCAAAGCGCGGGAGAGCGGTGTTGCTATTGGCGCG 2940
2941 TCTTCGGTCTCCTCGCTCACTGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 3000
2941 TCTTCGGTCTCCTCGCTCACTGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 3000
3001 TCAGCTCACTCAAGCGGTAAATACGGTTATCCAGAAATCGAGCTCAAGTCAGAGG 3060
3001 TCAGCTCACTCAAGCGGTAAATACGGTTATCCAGAAATCGAGCTCAAGTCAGAGG 3060
3061 AACATGTGAGCAAAAGCGCCAGCAAAAGCCGCTAAAGAGCCGCTGCTGCGG 3120
3061 AACATGTGAGCAAAAGCGCCAGCAAAAGCCGCTAAAGAGCCGCTGCTGCGG 3120
3121 TTTTTCATAGGCTCGCCCGCTCGAGAGCATCAAAATTCGACGCTCAAGTCAGAGG 3180
3121 TTTTTCATAGGCTCGCCCGCTCGAGAGCATCAAAATTCGACGCTCAAGTCAGAGG 3180
3181 TGGCGAAACCCGACAGGACTATAAGATACAGGGGTTCCTCCCTGGAAGCTCCCTCGTG 3240
3181 TGGCGAAACCCGACAGGACTATAAGATACAGGGGTTCCTCCCTGGAAGCTCCCTCGTG 3240
3241 CGCTCTCTGTTCCGACCCCTGCGGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGA 3300
3241 CGCTCTCTGTTCCGACCCCTGCGGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGA 3300
3301 AGGCTGGCGCTTCTCATAGCTCAAGCTGTAGGTATCTCAGTTGCTGCTGCTGCTGCTGCT 3360
3301 AGGCTGGCGCTTCTCATAGCTCAAGCTGTAGGTATCTCAGTTGCTGCTGCTGCTGCTGCT 3360
3361 TCCAAAGTGGCTGTGTGCAAGAACCCCGCTTACCGGACCGCTGCGCTTATCCGCT 3420
3361 TCCAAAGTGGCTGTGTGCAAGAACCCCGCTTACCGGACCGCTGCGCTTATCCGCT 3420
3421 AACTATCTGCTTTCAGTCCAAACCGGTAAAGACAGCTTATCGCCTGCGCAGCAGCACT 3480
3421 AACTATCTGCTTTCAGTCCAAACCGGTAAAGACAGCTTATCGCCTGCGCAGCAGCACT 3480
3481 GGTAAACAGGATTAAGCAGAGCGGTATGAGGGGTGCTACAGAGTCTTGAAGTGGTG 3540
3481 GGTAAACAGGATTAAGCAGAGCGGTATGAGGGGTGCTACAGAGTCTTGAAGTGGTG 3540
3541 CCTAACTACGGCTACACTAGAGAAAGTATTTGGTATCTGCGCTCTGCTGGAAGCCAGTT 3600
3541 CCTAACTACGGCTACACTAGAGAAAGTATTTGGTATCTGCGCTCTGCTGGAAGCCAGTT 3600
3601 ACCTTCGGAAGAGTGTGTAGTCTTGTATCGGCAAAACAAACCCAGCTGAGCGGT 3660
3601 ACCTTCGGAAGAGTGTGTAGTCTTGTATCGGCAAAACAAACCCAGCTGAGCGGT 3660
3661 GGTGTTTTGTTTTCAGAGCAGGATTAAGCAGAAAGAGGATCTCAAGAGATCCT 3720
3661 GGTGTTTTGTTTTCAGAGCAGGATTAAGCAGAAAGAGGATCTCAAGAGATCCT 3720
3721 TTGATCTTTTCTACGGGCTCTGACGCTCAGTGAAGAAAGGAACTCAGCTTAAAGGATTTG 3780
3721 TTGATCTTTTCTACGGGCTCTGACGCTCAGTGAAGAAAGGAACTCAGCTTAAAGGATTTG 3780
3781 GTCATGAGATTAAGTCGACAAAGCGGCTCAGTGGCTCCCTCCCTCCTCAGTTCGGGG 3840
3781 GTCATGAGATTAAGTCGACAAAGCGGCTCAGTGGCTCCCTCCCTCCTCAGTTCGGGG 3840
3841 GCATGGATCGCGATAGCGCTGCTGTTTCTGATGCGGACCGGATTTGCACTTCGCGG 3900
3841 GCATGGATCGCGATAGCGCTGCTGTTTCTGATGCGGACCGGATTTGCACTTCGCGG 3900
3901 TAGAACTCCGAGGCTCGTCCAGCTCAGGACGAGCTGAACCACTCGCGAGGGATTCG 3960

SQ Sequence 8349 BP; 1959 A; 2355 C; 2088 G; 1947 T; 0 other;									
Query Match 42.4%; Score 2240; DB 21; Length 8349;									
Best Local Similarity 93.4%; Pred. No. 0;									
Matches 2409; Conservative 0; Mismatches 25; Indels 144; Gaps 1;									
QY	2706	CTAGAGCTAATCATGTCATAGCTGTTTCTGTCGAAATGTTATCCGCTCACAAATCC	2765						
DB	5911	CTTGGGGTAAATCATGTGTCATAGCTGTTTCTGTCGAAATGTTATCCGCTCACAAATCC	5970						
QY	2766	ACACAATACAGAGCGGAGCATAAAGTCTAAAGCTCGGGTGCCTAATGATGAGCTA	2825						
DB	5971	ACACAATACAGAGCGGAGCATAAAGTCTAAAGCTCGGGTGCCTAATGATGAGCTA	6030						
QY	2826	ACTCACATTAATTTGCGTTGCGCTCACTCGCCCGTTTTCCAGTCGGGAAACCTCTCGTGCCA	2885						
DB	6031	ACTCACATTAATTTGCGTTGCGCTCACTCGCCCGTTTTCCAGTCGGGAAACCTCTCGTGCCA	6090						
QY	2886	GCTGCATTAAATGAAATCGGCCCAACGCGGGGAGAGCGGTTTTCGTTATTTGGCGCTCTTC	2945						
DB	6091	GCTGCATTAAATGAAATCGGCCCAACGCGGGGAGAGCGGTTTTCGTTATTTGGCGCTCTTC	6150						
QY	2946	CGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT	3005						
DB	6151	CGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT	6210						
QY	3006	TCATCAAGCGGTAATACGGTTATCCACAGAAATCAGGGGTAACGACAGGAAAGAAAT	3065						
DB	6211	TCATCAAGCGGTAATACGGTTATCCACAGAAATCAGGGGTAACGACAGGAAAGAAAT	6270						
QY	3066	GTGAGCAAAAGGCGACGAAAGGCGGAGGACCGTAAAGGCGGCTGCTCGCGCTTTT	3125						
DB	6271	GTGAGCAAAAGGCGACGAAAGGCGGAGGACCGTAAAGGCGGCTGCTCGCGCTTTT	6330						
QY	3126	CCATAGGCTCCGCCCTTACAGAGATCAACAAATTCAGCGCTCAAGTCAGAGGTGGCG	3185						
DB	6331	CCATAGGCTCCGCCCTTACAGAGATCAACAAATTCAGCGCTCAAGTCAGAGGTGGCG	6390						
QY	3186	AAACCCGACAGGACTAAGATACAGAGGTTTCCCGCTGAGAGCTCCCTCGTGGGCTC	3245						
DB	6391	AAACCCGACAGGACTAAGATACAGAGGTTTCCCGCTGAGAGCTCCCTCGTGGGCTC	6450						
QY	3246	TCCTGTTCCAGCCCTGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAGCGT	3305						
DB	6451	TCCTGTTCCAGCCCTGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAGCGT	6510						
QY	3306	GGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGTTGAGGTGCTTCGCTCCAA	3365						
DB	6511	GGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGTTGAGGTGCTTCGCTCCAA	6570						
QY	3366	GCTGGGCTGTGCAGAACCCCGCTTCAGCCCGACCGCTGCGCTTATCCGTTAACTA	3425						
DB	6571	GCTGGGCTGTGCAGAACCCCGCTTCAGCCCGACCGCTGCGCTTATCCGTTAACTA	6630						
QY	3426	TCGCTTTAGTCCAAACCCCGTAAAGACAGCTTATCCGCACTGGCAGCCACTGGTAA	3485						
DB	6631	TCGCTTTAGTCCAAACCCCGTAAAGACAGCTTATCCGCACTGGCAGCCACTGGTAA	6690						
QY	3486	CAGGATTAGCAGCGAGGTATAGCGGTGTACAGAGTTCTTGAAGTGTGGCTTAA	3545						
DB	6691	CAGGATTAGCAGCGAGGTATAGCGGTGTACAGAGTTCTTGAAGTGTGGCTTAA	6750						
QY	3546	CTACGGCTACACTAGAGAACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTT	3605						
DB	6751	CTACGGCTACACTAGAGAACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTT	6810						
QY	3606	CGGAAAAGAGTTGGTAGCTCTTGTATCCGGCAAAACAAACGACGCTGGTAGCGGTGTT	3665						
DB	6811	CGGAAAAGAGTTGGTAGCTCTTGTATCCGGCAAAACAAACGACGCTGGTAGCGGTGTT	6870						
QY	3666	TTTGTGTTGACGACGAGATACGGCGAGAAAAAAGGATCTCAAGAGAGATCCCTTTGAT	3725						
DB	6871	TTTGTGTTGACGACGAGATACGGCGAGAAAAAAGGATCTCAAGAGAGATCCCTTTGAT	6930						

QY	3726	CTTTTCTACGGGCTCAGCGCTCAGTGGAAACGAAACTCAGTTTAAGGGATTTTGGTCAT	3785						
DB	6931	CTTTTCTACGGGCTCAGCGCTCAGTGGAAACGAAACTCAGTTTAAGGGATTTTGGTCAT	6990						
QY	3786	GAGATTATCGTCGACAAAGCGGCATCGTGCCTCCCACTCTCTGCAAGTTCGGGGGCATG	3845						
DB	6991	GAGATTATCAAAAAGGATCTTCACTAGATCCTT-----	7024						
QY	3846	GATCGCGGATAGCGCGTCTGCTGTTTCTCGATGCGGACGGAATTGCACTGCGCGGTAGAA	3905						
DB	7025	-----	7024						
QY	3906	CTCCGCGAGGCTGCTCCAGCCTCAGCGACAGCTGAACAACTCGCGAGGGATCGAGCCC	3965						
DB	7026	-----	7026						
QY	3966	GGGTGGCGGAGAACTCCAGCATGAGATCCCGCGCTGGAGGATCATCAGCGCGGCTC	4025						
DB	7027	GGGTGGCGGAGAACTCCAGCATGAGATCCCGCGCTGGAGGATCATCAGCGCGGCTC	7086						
QY	4026	CCGAAAAACGATTCGGAAGCCCAACCTTTTCATAGAGCGCGGTGGAATCGAAATCTCG	4085						
DB	7087	CCGAAAAACGATTCGGAAGCCCAACCTTTTCATAGAGCGCGGTGGAATCGAAATCTCG	7146						
QY	4086	TGATGGCAGGTTGGGCGTTCGTTGCTCGTCAATTTGAAACCCAGAGTCCCGCTCAGAAG	4145						
DB	7147	TGATGGCAGGTTGGGCGTTCGTTGCTCGTCAATTTGAAACCCAGAGTCCCGCTCAGAAG	7206						
QY	4146	AACCTGCTCAGAAGCGCATAGAGGCGATCGCTGCGAATCGGAGCGGCGATACCGTTAA	4205						
DB	7207	AACCTGCTCAGAAGCGCATAGAGGCGATCGCTGCGAATCGGAGCGGCGATACCGTTAA	7266						
QY	4206	AGCAGGAGAGCGGTGAGCCCATTTCCGCCCAAGCTTTTCAGCAATATCACCGGTAGCC	4265						
DB	7267	AGCAGGAGAGCGGTGAGCCCATTTCCGCCCAAGCTTTTCAGCAATATCACCGGTAGCC	7326						
QY	4266	AACCTGATGCTTATAGCGTCCGCCACACACAGCGCGGCACAGTCGATGAATCCAGAA	4325						
DB	7327	AACCTGATGCTTATAGCGTCCGCCACACACAGCGCGGCACAGTCGATGAATCCAGAA	7386						
QY	4326	AAGCGGCAATTTTCCACCATGATATTCGCAAGCAGGCAATCGCATGGGTCAAGACAGA	4385						
DB	7387	AAGCGGCAATTTTCCACCATGATATTCGCAAGCAGGCAATCGCATGGGTCAAGACAGA	7446						
QY	4386	TCCTCGCGCTGGGCATCGCGGCTTTCAGCTGGCGAACAAGTTCGGCTGGCGGAGCCCC	4445						
DB	7447	TCCTCGCGCTGGGCATCGCGGCTTTCAGCTGGCGAACAAGTTCGGCTGGCGGAGCCCC	7506						
QY	4446	TGATGCTCTTCTCGTCAGATCATCTGATCGACAAGACCGGCTTTCATCCGAGTACGTGCT	4505						
DB	7507	TGATGCTCTTCTCGTCAGATCATCTGATCGACAAGACCGGCTTTCATCCGAGTACGTGCT	7566						
QY	4506	CGCTCGATGCGATGTTTTCGCTTGGTGTGGAATGGGCGAGTAGCCGATCAAGCGTATGC	4565						
DB	7567	CGCTCGATGCGATGTTTTCGCTTGGTGTGGAATGGGCGAGTAGCCGATCAAGCGTATGC	7626						
QY	4566	AGCCGCGCATTTGCATCAGCCATGATGATCTTCTCGCAGAGCAAGTCAGATGAC	4625						
DB	7627	AGCCGCGCATTTGCATCAGCCATGATGATCTTCTCGCAGAGCAAGTCAGATGAC	7686						
QY	4626	AGGAGATCTTCCCGCGCACTTCGCCCAATAGCAGCCAGTCCCTTCCCGTTTCAGTGAACA	4685						
DB	7687	AGGAGATCTTCCCGCGCACTTCGCCCAATAGCAGCCAGTCCCTTCCCGTTTCAGTGAACA	7746						
QY	4686	ACGTCGAGCAGCTGCGCAGGAAACCGCGTGTGGCGAGCCAGCATAGCCGGCTGCC	4745						
DB	7747	ACGTCGAGCAGCTGCGCAGGAAACCGCGTGTGGCGAGCCAGCATAGCCGGCTGCC	7806						
QY	4746	TCGTCCTTCGAGTTTCAATTCAGGGCACCGGACAGGTTCGCTTTTGACAAAAAAGAAACCGGGCGC	4805						
DB	7807	TCGTCCTTCGAGTTTCAATTCAGGGCACCGGACAGGTTCGCTTTTGACAAAAAAGAAACCGGGCGC	7866						

CC the origin of replication and replication stability sequences, and the
 CC promoter and coding sequence for t99 kanamycin/neomycin
 CC phosphotransferase. The plasmid was synthesised from 192 50-mers (see
 CC AAX52021-12) to demonstrate the method of the invention. The
 CC specification describes a method for the synthesis of replication
 CC competent double-stranded polynucleotides. The method comprises
 CC generating a first set of oligonucleotides corresponding to the plus
 CC strand and a second set corresponding to the minus strand and annealing.
 CC The method can be used for preparing polynucleotides encoding sequences
 CC involved in a biochemical pathway. In particular, they can be used to
 CC produce polynucleotides encoding enzymes, e.g. hexokinase, phosphohexose
 CC isomerase, phosphofructokinase-1, aldolase, triose-phosphate isomerase,
 CC glyceraldehyde-3-phosphate dehydrogenase, phosphoglycerate kinase,
 CC phosphoglycerate mutase, enolase or pyruvate kinase. They can
 CC also be used for the preparation of viral particles, artificial genomes
 CC and artificial genetic systems.

XX
 SQ Sequence 4800 BP; 1345 A; 1032 C; 1163 G; 1260 T; 0 other;

Query March 33.9%; Score 1793.4; DB 20; Length 4800;
 Best Local Similarity 91.8%; Pred. No. 4.9e-269;
 Matches 1980; Conservative 0; Mismatches 1; Indels 176; Gaps 1;

QY	3127	CATAGGCTCGCCCGCCCTGACGAGCATCAAAATCGAGCTCAAGTCAGAGTGGCGA	3186
DB	4800	CATAGGCTCGCCCGCCCTGACGAGCATCAAAATCGAGCTCAAGTCAGAGTGGCGA	4741
QY	3187	AACCCGACAGGACTATAAGATACAGCGCTTCCCTCGAGAGTCCCTCGTGGCTCT	3246
DB	4740	AACCCGACAGGACTATAAGATACAGCGCTTCCCTCGAGAGTCCCTCGTGGCTCT	4681
QY	3247	CTGTGTCGACGCTCGCGCTTACCGGATACCTGCGCGCTTCTCCCTCGGGAAGCGTG	3306
DB	4680	CTGTGTCGACGCTCGCGCTTACCGGATACCTGCGCGCTTCTCCCTCGGGAAGCGTG	4621
QY	3307	GGCTTTCTCATAGCTACGCTGTAGGTATCTCAGTTCGGTGTAGTCTGCTCGAAG	3366
DB	4620	GGCTTTCTCATAGCTACGCTGTAGGTATCTCAGTTCGGTGTAGTCTGCTCGAAG	4561
QY	3367	CTGGGCTGTGTGACGAAACCCCGCTTCAGCGCCGACCGCTGCGCTTATCCGGTAACTAT	3426
DB	4560	CTGGGCTGTGTGACGAAACCCCGCTTCAGCGCCGACCGCTGCGCTTATCCGGTAACTAT	4501
QY	3427	CGTCTTGAGTCCAAACCCGCTAGACGACACTTTCGCTACCTGCGCAGCAGCTGCTAC	3486
DB	4500	CGTCTTGAGTCCAAACCCGCTAGACGACACTTTCGCTACCTGCGCAGCAGCTGCTAC	4441
QY	3487	AGGATTAGCAGACGAGGTATGAGCGGTCTACAGAGTCTTTGAAAGTGTGCGCTTAAC	3546
DB	4440	AGGATTAGCAGACGAGGTATGAGCGGTCTACAGAGTCTTTGAAAGTGTGCGCTTAAC	4381
QY	3547	TACGGCTACACTAGAGACGATTTTGGTATCTGCGCTCTGCTGAGCCAGTTCCTTC	3606
DB	4380	TACGGCTACACTAGAGACGATTTTGGTATCTGCGCTCTGCTGAGCCAGTTCCTTC	4321
QY	3607	GGAAAAGAGTGTGAGTCTTTGATCCGGCAAAACAAACACCGCTGTGAGCGGTGTTT	3666
DB	4320	GGAAAAGAGTGTGAGTCTTTGATCCGGCAAAACAAACACCGCTGTGAGCGGTGTTT	4261
QY	3667	TTTGTGTAAGCAGCAGATTAACGCGCAGAAAAGGATCTCAAGAGATCTTTTGATC	3726
DB	4260	TTTGTGTAAGCAGCAGATTAACGCGCAGAAAAGGATCTCAAGAGATCTTTTGATC	4201
QY	3727	TTTCTACGGGTCTGACGCTCAGTGGAAACGAAACTCAGTTAAGGGATTTTGGTCATG	3786
DB	4200	TTTCTACGGGTCTGACGCTCAGTGGAAACGAAACTCAGTTAAGGGATTTTGGTCAT	4142
QY	3787	AGATTATCTCGACAAACGCGCCATCGTGCTCTCCCACTCTCTGCTGGCGGCGATGG	3846
DB	4141	-----	4142
QY	3847	ATCGCGGATAGCGCTGTGTTTCTCTGGATGCGGACGGATTTGCACTCCCGGTAGAAC	3906

DB	4141	-----	4142
QY	3907	TCCGCGAGTGTGTCAGCCTCAGGCAGCAGCTGAACCACTCGCGAGGGATCGAGCCCG	3966
DB	4141	-----	4137
QY	3967	GGTGGCGGAGNACTCCAGCATGAGATCCCGCGCTCGAGGATCATCCAGCGCGCTCC	4026
DB	4136	GGTGGCGGAGNACTCCAGCATGAGATCCCGCGCTCGAGGATCATCCAGCGCGCTCC	4077
QY	4027	CGAAAACGATTCCGAAGCCCAACCTTTATAGAAGCGGGGTGAATCGAAATCTCGT	4086
DB	4076	CGAAAACGATTCCGAAGCCCAACCTTTATAGAAGCGGGGTGAATCGAAATCTCGT	4017
QY	4087	GATGGCAGGTTGGCGCTGCTTGGTCTGCTCATTTGAAACCCAGAGTCCCGCTCAGAAGA	4146
DB	4016	GATGGCAGGTTGGCGCTGCTTGGTCTGCTCATTTGAAACCCAGAGTCCCGCTCAGAAGA	3957
QY	4147	ACTGCTCAAGAAGCGCATAGAAGCGATGCGTTCGGAATCGGAGCGCGGATACCGTAAA	4206
DB	3956	ACTGCTCAAGAAGCGCATAGAAGCGATGCGTTCGGAATCGGAGCGCGGATACCGTAAA	3897
QY	4207	GCACGAGGAAGCGGTGAGCCCATTTCCCGCCCAAGCTCTTCAGCAATATCAAGGATAGCA	4266
DB	3896	GCACGAGGAAGCGGTGAGCCCATTTCCCGCCCAAGCTCTTCAGCAATATCAAGGATAGCA	3837
QY	4267	ACGTATATGCTGATAGCGGTGCGCACACCCAGCGCGGCGACAGTCGATGAATCCAGAAA	4326
DB	3836	ACGTATATGCTGATAGCGGTGCGCACACCCAGCGCGGCGACAGTCGATGAATCCAGAAA	3777
QY	4327	AGGGGCCATTTTCCACCATGATTTTCGGCAAGCAGGATTCGCCATGGGTACACGAGAT	4386
DB	3776	AGGGGCCATTTTCCACCATGATTTTCGGCAAGCAGGATTCGCCATGGGTACACGAGAT	3717
QY	4387	CTTCGCGCTCGGCGATCGCGCTTGAGCTGGCGCAACAGTTCGCTGCGCGCGAGCCCT	4446
DB	3716	CTTCGCGCTCGGCGATCGCGCTTGAGCTGGCGCAACAGTTCGCTGCGCGAGCCCT	3657
QY	4447	GATGCTCTTCGCTCCAGATCATCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGTC	4506
DB	3656	GATGCTCTTCGCTCCAGATCATCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGTC	3597
QY	4507	GCTCGATGCGATGTTTCGGTGTGATGCGAATGGCAGGTAGCGGATCAAGCGTATGCA	4566
DB	3596	GCTCGATGCGATGTTTCGGTGTGATGCGAATGGCAGGTAGCGGATCAAGCGTATGCA	3537
QY	4567	GCAGCGCATTTGATCAGCATGATGATGATCTTCGGCAGGAGCAAGGTGAGATGACA	4626
DB	3536	GCAGCGCATTTGATCAGCATGATGATGATCTTCGGCAGGAGCAAGGTGAGATGACA	3477
QY	4627	GGAGATCTTCGCGCGGCACTTCGCGCAATAGCAGCGAGTTCCTTCGCGCTTCAGTGCAA	4686
DB	3476	GGAGATCTTCGCGCGGCACTTCGCGCAATAGCAGCGAGTTCCTTCGCGCTTCAGTGCAA	3417
QY	4687	GTTCGAGCAGCTGTGGCGAAGAACGCGCGTGTGGCCAGCCAGATAGCGCGCTGCT	4746
DB	3416	GTTCGAGCAGCTGTGGCGAAGAACGCGCGTGTGGCCAGCCAGATAGCGCGCTGCT	3357
QY	4747	CGTCTTCGAGTTCATTCAGGGCAGCCGACAGGTTCGCTTCGACAAAAGAACCCGGGCGCC	4806
DB	3356	CGTCTTCGAGTTCATTCAGGGCAGCCGACAGGTTCGCTTCGACAAAAGAACCCGGGCGCC	3297
QY	4807	CTTCGCGTGAAGCGGAAACGCGCGTGTGGCCAGCCAGATAGCGCGCTGCT	4866
DB	3296	CTTCGCGTGAAGCGGAAACGCGCGTGTGGCCAGCCAGATAGCGCGCTGCT	3237
QY	4867	CATAGCGAATAGCTCTCCACCAAGCGCGCGGAGAACCTTCGCTGCAATCAATCTGTT	4926
DB	3236	CATAGCGAATAGCTCTCCACCAAGCGCGCGGAGAACCTTCGCTGCAATCAATCTGTT	3177
QY	4927	CAATCATGGGAGACGATCTCATCTCTGCTCTTGTATCAGATCTTGTATCCCTCGCGCATC	4986
DB	3176	CAATCATGGGAGACGATCTCATCTCTGCTCTTGTATCAGATCTTGTATCCCTCGCGCATC	3117

QY 4987 AGATCTTGGCGGCAAGAAAGCCATCCAGTTACTTTGAGGGCTTCCCAACTTACCAG 5046
 Db 3116 AGATCTTGGCGGCAAGAAAGCCATCCAGTTACTTTGAGGGCTTCCCAACTTACCAG 3057
 QY 5047 AGGGCGCCAGCTGGCAATTCGGTTTGGCTTGTCTGCTGCTCCATAAACCCGCAAGTCTAGCT 5106
 Db 3056 AGGGCGCCAGCTGGCAATTCGGTTTGGCTTGTCTGCTGCTCCATAAACCCGCAAGTCTAGCT 2997
 QY 5107 ATCCGCATGTAAGCCCACTGCAAGCTACCTGCTTCTTCTTGGCTTGGCTTTCCTCCCTG 5166
 Db 2996 ATCCGCATGTAAGCCCACTGCAAGCTACCTGCTTCTTCTTGGCTTGGCTTTCCTCCCTG 2937
 QY 5167 TCCAGATAGCCCACTGCAAGCTACCTGCTTCTTCTTGGCTTGGCTTTCCTCCCTG 5226
 Db 2936 TCCAGATAGCCCACTGCAAGCTACCTGCTTCTTCTTGGCTTGGCTTTCCTCCCTG 2877
 QY 5227 TCTAGCTGTTCCGCTTCTTCTTGGCTTGGCTTTCCTCCCTG 5283
 Db 2876 TCTAGCTGTTCCGCTTCTTCTTGGCTTGGCTTTCCTCCCTG 2820

RESULT 5

AAV33629/c

ID AAV33629 standard; DNA; 5594 BP.

XX AAV33629;

DT 29-DEC-1998 (first entry)

DE GENSA 981, a monomeric DNA sequence produced by the invention.

KW Class IIS restriction endonuclease recognition site;

KW endogenous mouse promoter element; tissue-specific gene expression;

KW hormone-specific gene expression; as;

KW developmental-specific gene expression.

OS Synthetic.

XX WO9838326-A1.

XX 03-SEP-1998.

XX 28-FEB-1998; 58WO-US03918.

XX 28-FEB-1997; 97US-0070910.

XX (NATU-) NATURE TECHNOLOGY CORP.

XX Hodgson CP, Xu G, Zink NA;

XX WPI; 1998-495399/42.

XX Method for assembling gene or gene vector - comprises use of primers containing

XX Example 2; Pages 107-110; 141pp; English.

XX The invention provides a novel method for directing self-assembly of a gene having three or more fragments in a directionally and spatially ordered fashion to produce a gene or a gene vector. The method involves usage of primers, containing class IIS restriction endonuclease recognition sites, for isolation of these fragments. As described in the disclosure, the method may also use a vector for the incorporation and screening of endogenous mouse promoter elements for the identification of cell specific promoters. In the example given, plasmids pBK-CMV (AAV33626), pVLMs (AAV33623) and pVLOvH-900 (AAV33621) were used as templates from which six fragments were amplified. Each of the fragments contained different regulatory sequences. The six PCR fragments were designed to self-assemble into a retro-vector using the method of the invention. The present sequence, designated as GENSA 981, represents the monomeric DNA sequence of the six ligated fragments. In general, the method is claimed to be useful for

CC isolating and identifying regulatory sequences from a cell, including CC those for enhanced biological activity, or tissue-specific, CC hormone-specific or developmental-specific gene expression.

XX Sequence 5594 BP; 1232 A; 1519 C; 1518 G; 1325 T; 0 other;

QY Query Match 31.3%; Score 1652.4; DB 19; Length 5594;

Db Best Local Similarity 85.4%; Pred. No. 3.2e-247;

QY Matches 1916; Conservative 0; Mismatches 261; Indels 67; Gaps 4;

QY 3107 CGCGCTTCTGCTGGCTTTTCCATAGGCTCCGCCCTTACAGGATACCAAAATCGAC 3166

Db 5591 CGCGCTTCTGCTGGCTTTTCCATAGGCTCCGCCCTTACAGGATACCAAAATCGAC 5532

QY 3167 GCTCAAGTCAGAGTGGCGAACCAGCAGAGCTATAAAGATACCAAGGCTTTCCCTCG 3226

Db 5531 GCTCAAGTCAGAGTGGCGAACCAGCAGAGCTATAAAGATACCAAGGCTTTCCCTCG 5472

QY 3227 GAAGTCTCCCTGCTGGCTTCTCTGTTCCGACCCCTTACCGGATACCTCTCCGCT 3286

Db 5471 GAAGTCTCCCTGCTGGCTTCTCTGTTCCGACCCCTTACCGGATACCTCTCCGCT 5412

QY 3287 TTCTCCCTTCCGGAAGCGTGGCTTTCTCATAGCTCACGCTGTAGTATCTCAGTTCGG 3346

Db 5411 TTCTCCCTTCCGGAAGCGTGGCTTTCTCATAGCTCACGCTGTAGTATCTCAGTTCGG 5352

QY 3347 TCTAGTCTGCTTCCAGCTGGCTGTGTGCAAGAACCCCGCTTACGCCCGAGCGCT 3406

Db 5351 TGTAGTCTGCTTCCAGCTGGCTGTGTGCAAGAACCCCGCTTACGCCCGAGCGCT 5292

QY 3407 GCGCTTATCCGCTAACTATCTGTTGAGTCCAAACCCGCTAAGACACAGCTTATCCGCT 3466

Db 5291 GCGCTTATCCGCTAACTATCTGTTGAGTCCAAACCCGCTAAGACACAGCTTATCCGCT 5232

QY 3467 TGGCAGCAGCCTGCTTAAAGATTTAGCAGCAGGATTTAGCGGCTTCTCAGAGT 3526

Db 5231 TGGCAGCAGCCTGCTTAAAGATTTAGCAGCAGGATTTAGCGGCTTCTCAGAGT 5172

QY 3527 TCTTGAAGTGGTGGCTTAACTACGCTTACACTAGAAGAACAGTATTTGGTATCTGGGCTC 3586

Db 5171 TCTTGAAGTGGTGGCTTAACTACGCTTACACTAGAAGAACAGTATTTGGTATCTGGGCTC 5112

QY 3587 TGTCAAGCCAGTTACCTTCGGAAGAGTTGGTACTCTTGTATCGGCAACCAACCA 3646

Db 5111 TGTCAAGCCAGTTACCTTCGGAAGAGTTGGTACTCTTGTATCGGCAACCAACCA 5052

QY 3647 CCGCTGGTAGCGGTGTTTCTTTTCTTCAAGCAGCAGATTTACGCCGAGAAAGGAT 3706

Db 5051 CCGCTGGTAGCGGTGTTTCTTTTCTTCAAGCAGCAGATTTACGCCGAGAAAGGAT 4992

QY 3707 CTCAAGAGATCCCTTGTATCTTCTACGGGTCTGACGCTC-----AGTGGAGC 3757

Db 4991 GGTGTGGCTCTTTTATTGAGCTCGGGAGCAGAACCGCGCAACAGAGCAGAGCG 4932

QY 3758 AAAACTCAGTTAAGGATTTTGTCTATGATTTATCTGACACAAAGCGGCTCATCGTGC 3817

Db 4931 AACTGATTTGTTAGTTCAATAAGGCACAGGTCATTTTCTAGGTCCTTGGGCGACCTTGA 4872

QY 3818 CTC-----CCACTCTCTGAGTTTCGGGGCATGATCGCGGATAGCGGCTG 3864

Db 4871 AACTCTGATGTTCTCTAGAAATGCTGAGGCTGACCGCATCTGGGAGCATCTGTT 4812

QY 3865 CTGGTTTCTGATGCCGACCGATTTG-----ATGCGGTAGAACTCCGCGAG 3914

Db 4811 CTGGGCTTGGAGCGGGGCGAGAACTGCTTACCACAGATATCTGTTTGGCCCATATTC 4752

QY 3915 GTCTGTCAGCTCAGGAGCAGCTGAACCACTCGCAGGGGATTCGAGCCCGGGGTGGC 3974

Db 4751 GTCTGTCATCTGTTCTTGGCCCTGAGCGGGGAGGAACTGCTTACCACAGATATCTG 4692

QY 3975 GAAGAACTCCAGCATGATATCCCGCTGGAGGATCATCCAGCCCGGCTCCCGGAAAC 4034

Db 4691 TTGGCCCATATTCAGCTGTTTCCATCTCTTCTGACCTTGTGTAACCTTCTTATTTCT 4632

QY 4035 GATTCGAGCCCAACCTTTTCATAGAGCGCGGTGGAATCGAATCTCG----- 4085
DB 4631 AGTATGATATTTTCCATGCTTGCATAAATGGGTACTTAACTAGCTTGCCAACTTA 4572
QY 4086 -----TGATGGCAGGTTGGCGTCTGGTGGTGGTCACTT 4119
DB 4571 CAGTGGGGTCTTCAATCCGCCCTTTTCTGGAGGTTGGCGTCTGGTGGTCACTT 4512
QY 4120 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGCGATAGAGGCGATCGCT 4179
DB 4511 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGCGATAGAGGCGATCGCT 4452
QY 4180 GCGATCCGGGCGCGATACCGTAACAGCAGGAGCGGTGAGCCCATTCGCGCCCAA 4239
DB 4451 GCGAATCCGGAGCGCGATACCGTAACAGCAGGAGCGGTGAGCCCATTCGCGCCCAA 4392
QY 4240 GCTCTTCAGCAATATCACGGTAGCCAAAGCTATGCTCTGATAGCGGTGCGCCACACCA 4299
DB 4391 GCTCTTCAGCAATATCACGGTAGCCAAAGCTATGCTCTGATAGCGGTGCGCCACACCA 4332
QY 4300 GCGGCGCCAGTCCGATGATCCAGAAAGCGGCGCATTTTCCACCATGATATTCGCGAAGC 4359
DB 4331 GCGGCGCCAGTCCGATGATCCAGAAAGCGGCGCATTTTCCACCATGATATTCGCGAAGC 4272
QY 4360 AGGCATCCGATCGGTTCAGCAGAGATCTCGCGTCCGGGATCGCGCGCTTGAGCTGG 4419
DB 4271 AGGCATCCGATCGGTTCAGCAGAGATCTCGCGTCCGGGATCGCGCGCTTGAGCTGG 4212
QY 4420 CGAACAGTTCCGCTCGCGCGAGCCCTGATGCTCTTCTCCAGATCATCTGATGACAA 4479
DB 4211 CGAACAGTTCCGCTCGCGCGAGCCCTGATGCTCTTCTCCAGATCATCTGATGACAA 4152
QY 4480 GACCGCTTCATCCGATGATGCTGCTCGATGCGATGTTTCGCTTGGTGGTGAATG 4539
DB 4151 GACCGCTTCATCCGATGATGCTGCTCGATGCGATGTTTCGCTTGGTGGTGAATG 4092
QY 4540 GCGAGTCCGATCAAGCGTATGCGAGCGCGCGCATTTGATCAGCCATGATGATCTT 4599
DB 4091 GCGAGTCCGATCAAGCGTATGCGAGCGCGCGCATTTGATCAGCCATGATGATCTT 4032
QY 4600 TCTCGCAGGAGCAAGTGAAGTGAAGAGGATCTTCCCGCGGCACTTCCGCCAATAGCA 4659
DB 4031 TCTCGCAGGAGCAAGTGAAGTGAAGAGGATCTTCCCGCGGCACTTCCGCCAATAGCA 3972
QY 4660 GCGAGTCCCTTCCGCTTCAGTGAACAGTTCGAGCAGAGTTCGCGAGGAAACCGCGTCG 4719
DB 3971 GCGAGTCCCTTCCGCTTCAGTGAACAGTTCGAGCAGAGTTCGCGAGGAAACCGCGTCG 3912
QY 4720 TGGCGAGCAGATAGCGCGCTGCTGCTCGATGAGTTCATTCAGGCGACCGGACAGT 4779
DB 3911 TGGCGAGCAGATAGCGCGCTGCTGCTCGATGAGTTCATTCAGGCGACCGGACAGT 3952
QY 4780 CGGTCTTGACAAAAGAACCGGCGCGCTTCCGCTGACAGCGGAAACACCGCGGCACTAG 4839
DB 3851 CGGTCTTGACAAAAGAACCGGCGCGCTTCCGCTGACAGCGGAAACACCGCGGCACTAG 3792
QY 4840 AGCAGCGGATGCTGTTGTCGCGAGTCAATAGCGGATAGCTTCCACCGACGCGCGCG 4899
DB 3791 AGCAGCGGATGCTGTTGTCGCGAGTCAATAGCGGATAGCTTCCACCGACGCGCGCG 3732
QY 4900 GAGAACCTCGTGCAATCATCTTCTTCAATCATGCGAAGCGATCTTCATCTGCTCTT 4959
DB 3731 GAGAACCTCGTGCAATCATCTTCTTCAATCATGCGAAGCGATCTTCATCTGCTCTT 3672
QY 4960 GATCAGATCTGATCCCTGCGCATCAGATCTTGGCGGCAAGAACCATTCAGTTTA 5019
DB 3671 GATCAGATCTGATCCCTGCGCATCAGATCTTGGCGGCAAGAACCATTCAGTTTA 3612
QY 5020 CTTTGCAGGCTTCCCAACTTACAGAGGCGCGCCAGCTGCGCAATTCGGTTTCGCTTG 5079
DB 3611 CTTTGCAGGCTTCCCAACTTACAGAGGCGCGCCAGCTGCGCAATTCGGTTTCGCTTG 3552

QY 5080 CTCTCCATAAAACCGCCAGTCTAGCTATCGCATGTAAAGCCACTGCAAGCTACTCTCT 5139
DB 3551 CTCTCCATAAAACCGCCAGTCTAGCTATCGCATGTAAAGCCACTGCAAGCTACTCTCT 3492
QY 5140 TTCTCTTTCGCTTGGTTCCTTCCCTTGTCCAGATAGCCAGTAGCTGACATTCATCCGGG 5199
DB 3491 TTCTCTTTCGCTTGGTTCCTTCCCTTGTCCAGATAGCCAGTAGCTGACATTCATCCGGG 3432
QY 5200 GTCAGCACCGTTTCTCCGACTGGCTTCTAGCTTTCGCTTCTTTAGCAGCCCTTGC 5259
DB 3431 GTCAGCACCGTTTCTCCGACTGGCTTCTAGCTTTCGCTTCTTTAGCAGCCCTTGC 3372
QY 5260 GCCTGAGTCTTTCGCGCAGCGTG 5283
DB 3371 GCCTGAGTCTTTCGCGCAGCGTG 3348

RESULT 6

AAH74865/c

ID AAH74865 standard; DNA; 5285 BP.

XX AAH74865;

XX 29-OCT-2001 (first entry)

XX Nucleotide sequence of a construct comprising the PCV Rep gene.

XX Rolling circle replication; RCR; Rep gene; gene function; Gemminivirus;
XX Circovirus; Nanovirus; gene therapy; PCV; ss.

XX Synthetic.

XX Porcine circovirus.

XX WO200161024-A2.

XX 23-AUG-2001.

XX 15-FEB-2001; 2001WO-US05394.

XX 16-FEB-2000; 2000US-0505477.

XX (LARG-) LARGE SCALE BIOLOGY CORP.

XX Palmer KE, Pogue GP;

XX WPI; 2001-522601/57.

XX New polynucleotide capable of rolling circle replication in host.

XX useful for discovery of gene function, comprises Rep gene, sequences
XX that are cis on the polynucleotide and open reading frame encoding
XX protein of interest

XX Example 1; Fig 5; 61pp; English.

XX The specification describes a polynucleotide that is capable of rolling
XX circle replication (RCR) in an eukaryotic host. The polynucleotide
XX comprises a Rep gene encoding from a virus, sequences that are cis on
XX the polynucleotide such that the Rep protein can bring about RCR of the
XX polynucleotide, an open reading frame encoding a protein of interest
XX capable of being expressed in the host, and optionally a multiple
XX cloning site, where the polynucleotide lacks genes of the virus. The
XX virus is selected from the group of genera of family Gemminiviridae,
XX Circoviridae and Nanovirus. The polynucleotide sequence is useful for
XX the discovery of the function of genes in eukaryotic hosts, and for
XX inducing or enhancing a trait in a host eukaryotic cell, for
XX down-regulating a gene in a plant or in mammalian cell and thus altering
XX or even eliminating the function of that gene, as gene sequence delivery
XX tools for mammalian genomic approaches, for gene therapy applications
XX (for delivering therapeutic or complementing gene products to organisms
XX or cells), and in whole animal genomics. The present sequence represents
XX a construct comprising the whole PCV genome. This construct contains the
XX PCV Rep gene under the transcriptional control of its own promoter, and
XX has the putative coat protein inactivated by insertion of the bacterial

Db 3245 GAGCAGCTGCGCAGGAAACCCCGTGTGGCCAGCCACGATAGCCGCTGCTCCGTC 3186
 QY 4751 CTGAGTTTATTTCAGGACACCGGACAGTGTGGTCTTGACAAAAAGAACCGGGCGCCCTG 4810
 Db 3185 TTGAGTTTATTTCAGGACACCGGACAGTGTGGTCTTGACAAAAAGAACCGGGCGCCCTG 3126
 QY 4811 CGGTGACACCGGACACCGGCGGATCAGACAGCCGATGTCTGTGTGTCGCCAGTCATA 4870
 Db 3125 CGGTGACACCGGACACCGGCGGATCAGACAGCCGATGTCTGTGTGTCGCCAGTCATA 3066
 QY 4871 GCGAATAGCTCTCCACCCAGCGCGCGGAGAACCTGCGTGCAATCCATCTTGTTCAT 4930
 Db 3065 GCGAATAGCTCTCCACCCAGCGCGCGGAGAACCTGCGTGCAATCCATCTTGTTCAT 3006
 QY 4931 CATGCGAAACGATCTCTCATCTGCTCTTGATCATGATCTTGATCCCTGGCGCCATCAGAT 4990
 Db 3005 CATGCGAAACGATCTCTCATCTGCTCTTGATCATGATCTTGATCCCTGGCGCCATCAGAT 2946
 QY 4991 CTGTGGCGGCAAGAACGATCCATCTTTCAGGGCTTCCAACTTACAGAGG 5050
 Db 2945 CCTGTGGCGGCAAGAACGATCCATCTTTCAGGGCTTCCAACTTACAGAGG 2886
 QY 5051 CGCCCGAGCTGCAATTCGGTTCGCTGTGTCATATAAACCGCCAGCTAGCTATCG 5110
 Db 2885 CGCCCGAGCTGCAATTCGGTTCGCTGTGTCATATAAACCGCCAGCTAGCTATCG 2826
 QY 5111 CATGTAGCCCACTGCAAGTACTGCTTTCTCTTTGCGGTGCGTTTCCCTTGCCA 5170
 Db 2825 CCATGTAGCCCACTGCAAGTACTGCTTTCTCTTTGCGGTGCGTTTCCCTTGCCA 2766
 QY 5171 GATAGCCCACTGCAAGTACTGCTTTCTCTTTGCGGTGCGTTTCCCTTGCCA 5230
 Db 2765 GATAGCCCACTGCAAGTACTGCTTTCTCTTTGCGGTGCGTTTCCCTTGCCA 2706
 QY 5231 CGTG 5234
 Db 2705 CGTG 2702

RESULT 7

AAV33630/c
 ID AAV33630 standard; DNA; 6561 BP.
 XX AAV33630;
 XX
 XX
 DT 29-DEC-1998 (first entry)
 XX
 DE Plasmid VLSNO2 DNA sequence used in the method of the invention.
 XX
 XX Class IIS restriction endonuclease recognition site; LTR;
 KW endogenous mouse promoter element; tissue-specific gene expression;
 KW hormone-specific gene expression; mouse VL30 genome;
 KW developmental-specific gene expression; long terminal repeat; ss.
 XX
 OS Synthetic.
 XX
 PN WO983826-A1.
 XX
 PD 03-SEP-1998.
 XX
 PF 28-FEB-1998; 98WO-US03918.
 XX
 PR 28-FEB-1997; 97US-0070910.
 XX
 XX (NATU-) NATURE TECHNOLOGY CORP.
 FA
 XX Hodgson CP, Xu G, Zink MA;
 FI
 XX WPI; 1998-495399/42.
 DR
 XX Method for assembling gene or gene vector - comprises use of primers
 PT containing class IIS restriction endonuclease recognition sites
 XX

Example 3; Pages 110-113; 141pp; English.

PS The invention provides a novel method for directing self-assembly of
 XX a gene having three or more fragments in a directionally and spatially
 CC ordered fashion to produce a gene or a gene vector. The method involves
 CC usage of primers, containing class IIS restriction endonuclease
 CC recognition sites, for isolation of these fragments. As described
 CC in the disclosure, the method may also use a vector for the incorporation
 CC and screening of endogenous mouse promoter elements for the
 CC identification of cell specific promoters. In the example given,
 CC plasmid VLSNO2, containing a circularly permuted mouse VL30 genome, was
 CC used in the method of the invention. In general, the method is claimed
 CC useful for isolating and identifying regulatory sequences from a cell,
 CC including those for enhanced biological activity, or tissue-specific,
 CC hormone-specific or developmental-specific gene expression.

SQ Sequence 6561 BP; 1629 A; 1593 C; 1696 G; 1643 T; 0 other;

Query Match 30.6%; Score 1618.2; DB 19; Length 6561;
 Best Local Similarity 86.7%; Pred. No. 6.1e-242;
 Matches 1911; Conservative 0; Mismatches 23; Indels 269; Gaps 1;
 QY 3081 GCAAAAGGCCAGAACCGGTAAAGAGCCGCGTGTGCGGCTTTTCCATAGGCTCCGCC 3140
 Db 2305 GCAGGTGTCTCGAGCGCGCATCTCGCGCGTGTGCGCGTTTTCATAGGCTCCGCC 2246
 QY 3141 CCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGTGGCGAAACCCGACAGACT 3200
 Db 2245 CCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGTGGCGAAACCCGACAGACT 2186
 QY 3201 ATAAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCT 3260
 Db 2185 ATAAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCT 2126
 QY 3261 GCGCTTACCGGATACCTGTCGCGCTTCTCCCTTGGGAAAGCGTGGCGCTTCTCATAG 3320
 Db 2125 GCGCTTACCGGATACCTGTCGCGCTTCTCCCTTGGGAAAGCGTGGCGCTTCTCATAG 2066
 QY 3321 CTCACGCTGTAGTATCTCAGTTCGCTGTAGTTCGCTTCCGCTCCAAGCTGGCGCTGTGCA 3380
 Db 2065 CTCACGCTGTAGTATCTCAGTTCGCTGTAGTTCGCTTCCGCTCCAAGCTGGCGCTGTGCA 2006
 QY 3381 CGAACCCCGCTTCAGCGCGACCGCTGCGCTTATCCGCTAATCTGCTTGTAGTCCAA 3440
 Db 2005 CGAACCCCGCTTCAGCGCGACCGCTGCGCTTATCCGCTAATCTGCTTGTAGTCCAA 1946
 QY 3441 CCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGCTAACAGGATTAGCAGC 3500
 Db 1945 CCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGCTAACAGGATTAGCAGC 1886
 QY 3501 GAGGTATGTAGCGGCTGCTACAGAGTCTTGAAGTGGTGGCTTAACCTACGCTACACTAG 3560
 Db 1885 GAGGTATGTAGCGGCTGCTACAGAGTCTTGAAGTGGTGGCTTAACCTACGCTACACTAG 1826
 QY 3561 AAGAACAGTATTGGTATCTGCGCTCTGCTGAAGCAGTTACCTTGGGAAAGAGTTGG 3620
 Db 1825 AAGAACAGTATTGGTATCTGCGCTCTGCTGAAGCAGTTACCTTGGGAAAGAGTTGG 1766
 QY 3621 TAGCTCTTATCCGCAAAACAAACACCGCTGTGAGCGTGGTGTGTTTGTGTTGCAAGCA 3680
 Db 1765 TAGCTCTTATCCGCAAAACAAACACCGCTGTGAGCGTGGTGTGTTTGTGTTGCAAGCA 1706
 QY 3681 GCAGATTACCGCAGAAAAAGGATCTCAAGAGATCTTGTGATCTTTTCTACGGGGTC 3740
 Db 1705 GCAGATTACG----- 1696
 QY 3741 TGACGCTCAGTGGAAAGAAAACTCAGTTTAAAGGATTTTGGTTCATGAGATTATCTGAC 3800
 Db 1695 ----- 1696
 QY 3801 CAAAGCGGCACTGTGCTCCCACTCTCTGCAATTCGGGGGCGATGGATGCGCGGATAGCC 3860

Db 1695 ----- 1696
Qy 3861 GCTGCTGTTTCTGGATGCCGAGATTGCACTGCCGTAGAACTCCGGAGGTGTC 3920
Db 1695 ----- 1696
Qy 3921 CAGCCTCAGCAGCAGCTGAACCACTCCGAGGGGATCGAGCCCGGGTGGGGAAGAA 3980
Db 1695 ----- AATCCGGGGTGGGGAAGAA 1675
Qy 3981 CTCAGCATGAGATCCCGCGCTGGAGAGATCATCGAGCGGGCTCCCGGAAAAAGATTCC 4040
Db 1674 CTCAGCATGAGATCCCGCGCTGGAGAGATCATCGAGCGGGCTCCCGGAAAAAGATTCC 1615
Qy 4041 GAAGCCCAACTTTCATAGAGAGCGCGGTGGAATCGAAATCTCGTGATGCGAGGTTGGG 4100
Db 1614 GAAGCCCAACTTTCATAGAGAGCGCGGTGGAATCGAAATCTCGTGATGCGAGGTTGGG 1555
Qy 4101 CGTCGCTTGTGCGTCAATTTTCGAAACCCAGAGTCCGCTCAGAGAACTCGTCAAGAAAG 4160
Db 1554 CGTCGCTTGTGCGTCAATTTTCGAAACCCAGAGTCCGCTCAGAGAACTCGTCAAGAAAG 1495
Qy 4161 CGATAGAGAGCGATGCGCTCGGAATCGGAGCGGGGATACCGTAAAGCAGGAAAGCGG 4220
Db 1494 CGATAGAGAGCGATGCGCTCGGAATCGGAGCGGGGATACCGTAAAGCAGGAAAGCGG 1435
Qy 4221 TCAGCCCATTTCCGGCGCAGCTCTTCAGCAATATCACGGGTAGCCAGCTATGTCCTGA 4280
Db 1434 TCAGCCCATTTCCGGCGCAGCTCTTCAGCAATATCACGGGTAGCCAGCTATGTCCTGA 1375
Qy 4281 TAGCGGTCCGCCACACCCAGCGGGCCACAGTCCGATGAATCCAGAAAGCGGCCATTTTC 4340
Db 1374 TAGCGGTCCGCCACACCCAGCGGGCCACAGTCCGATGAATCCAGAAAGCGGCCATTTTC 1315
Qy 4341 ACCATGATATTCGCGAGGAGCATCGCATCGGTTCAGCAGAGATCTCCGCTCCGGC 4400
Db 1314 ACCATGATATTCGCGAGGAGCATCGCATCGGTTCAGCAGAGATCTCCGCTCCGGC 1255
Qy 4401 ATGCGCCTTGTAGCTTCGCGAAGTTCGGTTCGGCGAGCCCTGATGCTCTTCGTC 4460
Db 1254 ATGCGCCTTGTAGCTTCGCGAAGTTCGGTTCGGCGAGCCCTGATGCTCTTCGTC 1195
Qy 4461 AGATCATCTGATCGACAGACCGGTTCATCCGAGTACGTCGCTCCGATCGATGT 4520
Db 1194 AGATCATCTGATCGACAGACCGGTTCATCCGAGTACGTCGCTCCGATCGATGT 1135
Qy 4521 TTGCTTGTGGTGGATGGCAGGTAGCGGATCAAGCGTATGACCGCGCGCATTTGA 4580
Db 1134 TTGCTTGTGGTGGATGGCAGGTAGCGGATCAAGCGTATGACCGCGCGCATTTGA 1075
Qy 4581 TCAGCCATGATGATATCTTCTCGGAGGAGCAAGTGAATGACAGAGATCTTCGCC 4640
Db 1074 TCAGCCATGATGATATCTTCTCGGAGGAGCAAGTGAATGACAGAGATCTTCGCC 1015
Qy 4641 GGCACTTCCGCAATAGCAGCCAGTCCCTTCGCGCTTCAGTGACAAAGTTCAGAGCAGCT 4700
Db 1014 GGCACTTCCGCAATAGCAGCCAGTCCCTTCGCGCTTCAGTGACAAAGTTCAGAGCAGCT 955
Qy 4701 GCGCAAGGAACCGCGCTGCGCAGCAGATAGCGCGCTGCTCTCTTCGAGTTCA 4760
Db 954 GCGCAAGGAACCGCGCTGCGCAGCAGATAGCGCGCTGCTCTCTTCGAGTTCA 895
Qy 4761 TTCAGGCGACGGAGCAGGTGCTTTCGCAAAAGAACCGGGCGCCCTGCGTGCACAGC 4820
Db 994 TTCAGGCGACGGAGCAGGTGCTTTCGCAAAAGAACCGGGCGCCCTGCGTGCACAGC 835
Qy 4821 CCGAACCGCGGATCAGAGCAGCGATTTGCTGTGTCGCGCTATAGCCGAAATAGC 4880
Db 834 CCGAACCGCGGATCAGAGCAGCGATTTGCTGTGTCGCGCTATAGCCGAAATAGC 775
Qy 4881 CTCTCCACCCAGCGCGGAGAACCTGCTGCAATCTTGTTCATCATCGGAAC 4940
Db 774 CTCTCCACCCAGCGCGGAGAACCTGCTGCAATCTTGTTCATCATCGGAAC 715

RESULT 8

ABQ78294

ID ABQ78294 standard; DNA; 2192 BP.

XX AC ABQ78294;

XX AC ABQ78294;

DT 05-NOV-2002 (first entry)

XX Nucleotide sequence of plasmid pVC0289.

XX Pig; growth hormone releasing hormone; GHRH; growth rate;
XX lean body mass; insulin-like Growth Factor-I; IGF-I; milk production;
XX feed efficiency; somatotroph; growth-related disorder;
XX hypopituitary dwarfism; meat production; egg production; ss.
OS Synthetic.

XX WO200261037-A2.

XX 08-AUG-2002.

XX 12-DEC-2001; 2001WO-US48726.

XX 12-DEC-2000; 2000US-255021P.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX (ADVI-) ADVISYS.

XX Schwartz RJ, Carpenter RH, Draghia-Akli R, Kern DR, Smith RG;

XX WPI; 2002-619237/66.

XX Improving or enhancing growth, lean body mass, milk production, feed
XX efficiency or insulin-like Growth Factor-I levels, comprises
XX introducing a vector encoding a growth hormone releasing hormone into
XX an animal before or during gestation -

XX Example 2; Page 111-112; 113pp; English.

XX The present sequence represents plasmid pVC0289, which was used to
XX create a vector carrying cDNA encoding pig growth hormone releasing
XX hormone (GHRH) cDNA. Nucleic acids encoding GHRH are used in the
XX method of the invention. The specification describes a method for
XX improving or enhancing characteristics e.g. growth, lean body mass,
XX insulin-like Growth Factor (IGF)-I levels, growth rate and milk
XX production in an offspring, and for delaying birth of an offspring.
XX The method comprises introducing a vector, encoding GHRH, into cells

Db 2012 GGCCTTTCTCATAGCTCACGCTGTAGTATCTCAGTTCCGTTAGTCTCGTTCCGCTCAA 2071
QY 3366 GCTGGCTGTGTGACGAACCCCGCTTCAGCCCGACCGCTCGCTTATCCGGTAACTA 3425
Db 2072 GCTGGCTGTGTGACGAACCCCGCTTCAGCCCGACCGCTCGCTTATCCGGTAACTA 2131
QY 3426 TCGTCTTAGTCCAAACCCGCTAGACAGACTATTCGCCACTTCGGCAGCAGCAGTGTAA 3485
Db 2132 TCGTCTTAGTCCAAACCCGCTAGACAGACTATTCGCCACTTCGGCAGCAGCAGTGTAA 2191
QY 3486 CAGGATTAGCAGACGCGAGGTATGTAGCGGTGTCTACAGAGTTCTTTGAAGTGTGTCCCTAA 3545
Db 2192 CAGGATTAGCAGACGCGAGGTATGTAGCGGTGTCTACAGAGTTCTTTGAAGTGTGTCCCTAA 2251
QY 3546 CTAGCGGTACACTAGAGAACAGTATTTGGTATCTCGCTCTGCTGAAGCAGGTTACCTT 3605
Db 2252 CTAGCGGTACACTAGAGAACAGTATTTGGTATCTCGCTCTGCTGAAGCAGGTTACCTT 2311
QY 3606 CGGAAAGAGTGTGTAGCTTTGATCCGGCAACAAACCAACCGCTGTGTAGCGGTGTTT 3665
Db 2312 CGGAAAGAGTGTGTAGCTTTGATCCGGCAACAAACCAACCGCTGTGTAGCGGTGTTT 2371
QY 3666 TTTTGTGTGCAAGCAGCAGATTTAGCGCAGAAAAAGGATCTCAAGAGATCTCTTTGAT 3725
Db 2372 TTTTGTGTGCAAGCAGCAGATTTAGCGCAGAAAAAGGATCTCAAGAGATCTCTTTGAT 2431
QY 3726 TTTTGTGTGCAAGCAGCAGATTTAGCGCAGAAAAAGGATCTCAAGAGATCTCTTTGAT 3785
Db 2432 TTTTGTGTGCAAGCAGCAGATTTAGCGCAGAAAAAGGATCTCAAGAGATCTCTTTGAT 2449
QY 3786 CAGATTATCTGTCGACAAAGCGGCTATCTGCTCCCTCCTCCTGCTGCTGCGGCGCATG 3845
Db 2450 ----- 2449
QY 3846 GATCGCGGATAGCGGCTGTGTTTCTGATGCGGACGAGTTTGCATCTGCGGCTAGAA 3905
Db 2450 ----- 2449
QY 3906 CTCCGCGAGGTCTCCAGCTTCAGGACAGCTGAACCACTCGCGAGGGGATCGAGCC 3965
Db 2450 ----- 2449
QY 3966 GGGGTGGCGAAGAACTCCAGCATGATCCCGCGCTGGAGGATCATCCAGCCGCGCTC 4025
Db 2450 ----- 2449
QY 4026 CCGGAAACGATTCCGAAGCCCAACCTTTATAGAAGCGCGCTGGAATCGAAATCTCG 4085
Db 2450 ----- 2449
QY 4086 TGATGCGAGTTGGCGCTCGCTTGTGCTCATTTGCAACCCAGAGTCCCGCTCAGAAG 4145
Db 2450 ----- CGCTCAGAAG 2459
QY 4146 AACTGCTCAAGAGCGATAGAGCGATGCGGTGCGATTCGGAGCGCGGATCCGTTAA 4205
Db 2460 AACTGCTCAAGAGCGATAGAGCGATGCGGTGCGATTCGGAGCGCGGATCCGTTAA 2519
QY 4206 AGCAGAGAGCGGTTCAGCCCATTCGCGCCCAAGCTCTTCAGCAATATACCGGTAGCC 4265
Db 2520 AGCAGAGAGCGGTTCAGCCCATTCGCGCCCAAGCTCTTCAGCAATATACCGGTAGCC 2579
QY 4266 AAGCTATGCTCTGATAGGTTGCGCACACCGCGCGCAGCTGATGATCCAGAA 4325
Db 2580 AAGCTATGCTCTGATAGGTTGCGCACACCGCGCGCAGCTGATGATCCAGAA 2639
QY 4326 AAGCGGCCATTTTCCACCATGATATTCGGCAAGCAGGATCGCCATGGGTACGACGAGA 4385
Db 2640 AAGCGGCCATTTTCCACCATGATATTCGGCAAGCAGGATCGCCATGGGTACGACGAGA 2699
QY 4386 TCTGCGCGCTCGGCGATCGCGCTTGAAGCTGCGGACAGTTGCGTGTGCGGAGCCCC 4445

Db 2700 TCCTCGCGCTCGGCGATGCGCGCTTGAGCCTGGCGAACAGTTTCGGCTGGCGGAGCCCC 2759
QY 4446 TGATGCTTTCTGCTCCAGATCATCTGATCGAACAGACCGGCTTCCATCCGAGTACGTGCT 4505
Db 2760 TGATGCTTTCTGCTCCAGATCATCTGATCGAACAGACCGGCTTCCATCCGAGTACGTGCT 2819
QY 4506 CGCTCGATGCGATCTTTGCTTGGTGTGGAATGGGACAGTAGCCGATCAAGCGTATGC 4565
Db 2820 CGCTCGATGCGATCTTTGCTTGGTGTGGAATGGGACAGTAGCCGATCAAGCGTATGC 2879
QY 4566 AGCCGCGCATTTGATCATGAGCCATGATGATCTTCTCGCAGCAGCAAGGTGATGATAC 4625
Db 2880 AGCCGCGCATTTGATCATGAGCCATGATGATCTTCTCGCAGCAGCAAGGTGATGATAC 2939
QY 4626 AGGAGATCTCTCCCGCGCACTTTGCGCCCAATAGCAGCAGTCTTCTTCCGCTTCAGTGACA 4685
Db 2940 AGGAGATCTCTCCCGCGCACTTTGCGCCCAATAGCAGCAGTCTTCTTCCGCTTCAGTGACA 2999
QY 4686 AGCTCGACACAGCTGCGCAGGAAACCGCGCTGCTGCGCAGCCACATAGCCGCTGCC 4745
Db 3000 AGCTCGACACAGCTGCGCAGGAAACCGCGCTGCTGCGCAGCCACATAGCCGCTGCC 3059
QY 4746 TCGTCTCTCAGTTCACTTCAGGCGCACCGGACAGCTTCGTCTTGACAAAAAGAACCGGCGC 4805
Db 3060 TCGTCTCTCAGTTCACTTCAGGCGCACCGGACAGCTTCGTCTTGACAAAAAGAACCGGCGC 3119
QY 4806 CCTCGGCTGACCGCGGAAACAGCGCGGCTCAGAGCAGCCGATCTCTGTTGTGCCAG 4865
Db 3120 CCTCGGCTGACCGCGGAAACAGCGCGGCTCAGAGCAGCCGATCTCTGTTGTGCCAG 3179
QY 4866 TCATAGCGAATAGCTTCTCCACCAAGCGCGGAGAACCTCGGTGCAATCCATCTTGT 4925
Db 3180 TCATAGCGAATAGCTTCTCCACCAAGCGCGGAGAACCTCGGTGCAATCCATCTTGT 3239
QY 4926 TCAATCATGCGAAGCAGTCTCATCTGCTCTTGATCAGATCTTGATCCCTTCGCGCAT 4985
Db 3240 TCAATCATGCGAAGCAGTCTCATCTGCTCTTGATCAGATCTTGATCCCTTCGCGCAT 3299
QY 4986 CAGATCTTTGGCGGCAAGAACCCATCGAGTTACTTTGAGGCTTCCCACTTACCA 5045
Db 3300 CAGATCTTTGGCGGCAAGAACCCATCGAGTTACTTTGAGGCTTCCCACTTACCA 3359
QY 5046 GAGGCGCGCCAGCTGCGAATTCGCGTTCGCTTCTGCTCCATAAAACCGCGCGCTAGC 5105
Db 3360 GAGGCGCGCCAGCTGCGAATTCGCGTTCGCTTCTGCTCCATAAAACCGCGCGCTAGC 3419
QY 5106 TATCGCATGTAAAGCCCACTGCAAGCTACCTGCTTCTCTT 5146
Db 3420 AACTGTGGGAAGGCGGATCGGTGCGGCTCTCTCGCTATT 3460

RESULT 10
AAV50428
ID AAV50428 standard; DNA; 3599 bp.
XX AAV50428;
AC AAV50428;
XX 26-OCT-1998 (first entry)
DT 26-OCT-1998 (first entry)
XX
DE Plasmid pIG0552 lower sequence nucleotides 1-3599.
XX Human; IGF-1; insulin-like growth factor 1; urinary incontinence;
KW gene therapy; neurotrophic factor; ss.
XX Synthetic.
OS Homo sapiens.
XX W0983529-A1.
FN 06-AUG-1998.
XX 04-FEB-1998; 98WO-US02051.
PF
XX

PR 04-FEB-1997; 97US-0036862.
 XX (GENE-) GENEMEDICINE INC.
 XX Coleman M;
 XX WPI; 1998-437184/37.
 XX Treatment of urinary incontinence - by delivering nucleic acid
 XX vector for expression of growth factor or neurotrophic factor in
 XX tissue(s)
 XX Disclosure; Page 105-107; 117pp; English.
 XX A method has been developed of treating urinary incontinence (UI) in
 XX mammals. The method comprises delivering a nucleic acid vector for the
 XX expression of a growth factor or neurotrophic factor in a tissue or
 XX tissues. The present sequence represents the actual lower sequence
 XX of plasmid pIG0552 nucleotides 1-3599 from the present invention. Due
 XX to the growth and stimulatory effects of growth factors and neurotrophic
 XX factors, introducing these factors to degenerated muscles in the
 XX urinary system can improve UI by enhancing both their integrity and
 XX neural innervation.
 XX Sequence 3599 BP; 733 A; 1098 C; 1042 G; 726 T; 0 other;
 SQ Query Match 30.4%; Score 1603.8; DB 19; Length 3599;
 Best Local Similarity 82.8%; Pred. No. 1.1e-239;
 Matches 2022; Conservative 0; Mismatches 27; Indels 392; Gaps 1;
 QY 2706 CTGACGTAATCATGTGTCTAGCTGTTCTCTGTGGAATTTGTTATCGCTCACAATTC 2765
 DB 56 CTTGGCGTAATCATGTGTCTAGCTGTTCTCTGTGGAATTTGTTATCGCTCACAATTC 115
 QY 2766 ACACACATACGAGCGCGGAGCATAAAGTGAAGCCTGGGGTGCCTAATGAGTAGCTA 2825
 DB 116 ACACACATACGAGCGCGGAGCATAAAGTGAAGCCTGGGGTGCCTAATGAGTAGCTA 175
 QY 2826 ACTCAGATTAATTCGTTGGCTCATCTGCCCTCCCTCCAGTCGGGAACCTGTGTGCCA 2885
 DB 176 ACTCAGATTAATTCGTTGGCTCATCTGCCCTCCCTCCAGTCGGGAACCTGTGTGCCA 235
 QY 2886 GCTGCAATTAATGATCGGCAACCGCGCGGAGAGCGGTTTGGGTATTTGGCGCTCTTC 2945
 DB 236 GCTGCAATTAATGATCGGCAACCGCGCGGAGAGCGGTTTGGGTATTTGGCGCTCTTC 295
 QY 2946 CGCTTCTCTCCTCAGTCACTCGCTCGCTCGCTCGCTCGCTCGGCGAGCGGTATCAGC 3005
 DB 296 CGCTTCTCTCCTCAGTCACTCGCTCGCTCGCTCGCTCGGCGAGCGGTATCAGC 355
 QY 3006 TCACTCAAGCGGTAATACGGTTATCCACAGAAATCAGGGGATACCGCAGAAAGACAT 3065
 DB 356 TCACTCAAGCGGTAATACGGTTATCCACAGAAATCAGGGGATACCGCAGAAAGACAT 415
 QY 3066 GTGAGCAAAAGGCGCAGCAAAAGCGCAGGAACCGTAAAGCGCGGTTGCTGGCGTTTTT 3125
 DB 416 GTGAGCAAAAGGCGCAGCAAAAGCGCAGGAACCGTAAAGCGCGGTTGCTGGCGTTTTT 475
 QY 3126 CCATAGGCTCGGCGGCTTCAAGCATCAAAAATCAGCAATCAGCTCAAGTCAAGGTGGCG 3185
 DB 476 CCATAGGCTCGGCGGCTTCAAGCATCAAAAATCAGCAATCAGCTCAAGTCAAGGTGGCG 535
 QY 3186 AAAACCCACAGGACTAAAGATACAGGCGTTTCCCTCGGAAGCTCCCTCGTGGCGTC 3245
 DB 536 AAAACCCACAGGACTAAAGATACAGGCGTTTCCCTCGGAAGCTCCCTCGTGGCGTC 595
 QY 3246 TCCGTGTTCCGACCTCGGCTTACCGATACCTGTCCGCTTTCCTCCCTCGGGAAGGCT 3305
 DB 596 TCCGTGTTCCGACCTCGGCTTACCGATACCTGTCCGCTTTCCTCCCTCGGGAAGGCT 655
 QY 3306 GCGCGTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGTTGAGGTGCTTCGCTCCAA 3365
 DB 656 GCGCGTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGTTGAGGTGCTTCGCTCCAA 715

QY 3366 GCTGGGCTGTGTGTCACGAACCCCGCTTCAGCCCGACCGCTGCGCTTATTCGGTAACCTA 3425
 DB 716 GCTGGGCTGTGTGTCACGAACCCCGCTTCAGCCCGACCGCTGCGCTTATTCGGTAACCTA 775
 QY 3426 TCGTCTTTAGTCCAAACCGGTAAAGACACACACTTATTCGCGCACTGGCAGCAGCACTGGTAA 3485
 DB 776 TCGTCTTTAGTCCAAACCGGTAAAGACACACACTTATTCGCGCACTGGCAGCAGCACTGGTAA 835
 QY 3486 CAGGATTAGCAGAGCGAGGTATGTAGGCGTCTCTACAGAGTCTTGAAGTGGTGGCTTAA 3545
 DB 836 CAGGATTAGCAGAGCGAGGTATGTAGGCGTCTCTACAGAGTCTTGAAGTGGTGGCTTAA 895
 QY 3546 CTACGGCTACACTAGAAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTT 3605
 DB 896 CTACGGCTACACTAGAAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTT 955
 QY 3606 CGGAAAAGAGTTGGTAGTCTTTGATCCGGCAAAACAAACACCGCTGTGTAGCGGTGGTTT 3665
 DB 956 CGGAAAAGAGTTGGTAGTCTTTGATCCGGCAAAACAAACACCGCTGTGTAGCGGTGGTTT 1015
 QY 3666 TTTTGTGTCAGCAGCAGATTAACGGCAGAAAGAAAGATCTCMAAGAGATCCCTTTGAT 3725
 DB 1016 TTTTGTGTCAGCAGCAGATTAACGGCAGAAAGAAAGATCTCMAAGAGATCCCTTTGAT 1075
 QY 3726 CTTTCTTACGGGCTCTGACGCTCAGTGGAAACGAAACCTCACTGTTAAGGATTTTGGTTCAT 3785
 DB 1076 CTTTCTTACGGGCTCTGACGCTCAGTGGAAACGAAACCTCACTGTTAAGGATTTTGGTTCAT 1093
 QY 3786 GAGATTAATCTGCAACCAAGCGGCCATCTGCTCTCCCACTCTCTGCACTTCGGGGGCGATG 3845
 DB 1094 GAGATTAATCTGCAACCAAGCGGCCATCTGCTCTCCCACTCTCTGCACTTCGGGGGCGATG 1093
 QY 3846 GATCGCGGATAGCCGCTGCTGTTTCTTGGATGCGGACGGATTTGCACTCCCGGTAGAA 3905
 DB 1094 GATCGCGGATAGCCGCTGCTGTTTCTTGGATGCGGACGGATTTGCACTCCCGGTAGAA 1093
 QY 3906 CTCCGCGAGTCTCCAGCCTCAGGCGAGCTGAACCACTCCGCGAGGGGATCGAGCC 3965
 DB 1094 CTCCGCGAGTCTCCAGCCTCAGGCGAGCTGAACCACTCCGCGAGGGGATCGAGCC 1093
 QY 3966 GGGGTGGCGAAGAAATCCAGCAGTAGATCCCGCGCTGGAGGATCATCCAGCCGCGGTC 4025
 DB 1094 GGGGTGGCGAAGAAATCCAGCAGTAGATCCCGCGCTGGAGGATCATCCAGCCGCGGTC 1093
 QY 4026 CCGGAAAACGATTCGGAAGCCCAACCTTTCATAGAAGCGCGGTGGAATCGAAATCTCG 4085
 DB 1094 CCGGAAAACGATTCGGAAGCCCAACCTTTCATAGAAGCGCGGTGGAATCGAAATCTCG 1093
 QY 4086 TGATGGCAGGTTGGGCGTCTGTTGGTCTGATTTTCGAACCCCGAGTCCCGCTCAGAG 4145
 DB 1094 TGATGGCAGGTTGGGCGTCTGTTGGTCTGATTTTCGAACCCCGAGTCCCGCTCAGAG 1103
 QY 4146 AACTCGTCAAGAAAGGCGATAGAGCGATGCGCTCGAATCGGAGCGCGGATACCGTAA 4205
 DB 1104 AACTCGTCAAGAAAGGCGATAGAGCGATGCGCTCGAATCGGAGCGCGGATACCGTAA 1163
 QY 4206 AGCAGGAGGAGCGGTACGCCATTCGCGCGCAAGCTCTTCAGCAATATCAGGAGTACC 4265
 DB 1164 AGCAGGAGGAGCGGTACGCCATTCGCGCGCAAGCTCTTCAGCAATATCAGGAGTACC 1223
 QY 4266 AACGCTATGCTCTGATAGCGGTCCGCCACACCCAGCGGCCACAGTCCGATGAATCAGAA 4325
 DB 1224 AACGCTATGCTCTGATAGCGGTCCGCCACACCCAGCGGCCACAGTCCGATGAATCAGAA 1283
 QY 4326 AAGCGGCAATTTCCACATGATATTTCGGCAAGCAGGCAATCGCCATGGTCAACAGAGA 4385
 DB 1284 AAGCGGCAATTTCCACATGATATTTCGGCAAGCAGGCAATCGCCATGGTCAACAGAGA 1343
 QY 4386 TCCTCGCGCTCGGCGATGCGCGCTTGAAGCTTGGCGAAGCAGTTTCGCTGCGCGAGCCCC 4445
 DB 1344 TCCTCGCGCTCGGCGATGCGCGCTTGAAGCTTGGCGAAGCAGTTTCGCTGCGCGAGCCCC 1403

QY 4446 TGATGCTCTTCTGTCAGATCATCTGATCGACAGACCGGCTTCCATCCGAGTACGTGCT 4505
 DB 1404 TGATGCTCTTCTGTCAGATCATCTGATCGACAGACCGGCTTCCATCCGAGTACGTGCT 1463
 QY 4505 CCGTCGATGCGATGTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4565
 DB 1464 CCGTCGATGCGATGTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1523
 QY 4566 AGCCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4625
 DB 1524 AGCCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1583
 QY 4626 AGGAGATCTCTGCGCGGACATCTGCGGACATGATGATGATGATGATGATGATGATGAT 4685
 DB 1584 AGGAGATCTCTGCGCGGACATCTGCGGACATGATGATGATGATGATGATGATGATGAT 1643
 QY 4686 AGCTGAGCAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4745
 DB 1644 AGCTGAGCAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1703
 QY 4746 TCGTCTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4805
 DB 1704 TCGTCTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1763
 QY 4806 CCGTGGCTGAGCGGCAACACCGCGGCGATCAGAGCAGCGGATGCTGCTGCTGCTGCTGCT 4865
 DB 1764 CCGTGGCTGAGCGGCAACACCGCGGCGATCAGAGCAGCGGATGCTGCTGCTGCTGCTGCT 1823
 QY 4866 TCATAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4925
 DB 1824 TCATAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1883
 QY 4926 TCATAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4985
 DB 1884 TCATAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1943
 QY 4986 CAGATCCTTGGCGGCAAGAAAGCCATCCAGTTACTTTGCGGGCTTCCCAACCTTACCA 5045
 DB 1944 CAGATCCTTGGCGGCAAGAAAGCCATCCAGTTACTTTGCGGGCTTCCCAACCTTACCA 5103
 QY 5046 GAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 5105
 DB 2004 GAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2063
 QY 5106 TATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5146
 DB 2064 AACTGTTGGGAGGCGGATCGGTCGGGCGCTCTCTGCTATT 2104

RESULT 11

AAV40796

AAV40796

AAV40796

AAV40796

AAV40796

AAV40796

AAV40796

AAV40796

AAV40796

AAV40796

AAV40796

AAV40796

AAV40796

AAV40796

AAV40796

AAV40796

AAV40796

AAV40796

AAV40796

AAV40796

AAV40796

AAV40796

AAV40796

AAV40796

AAV40796

AAV40796

AAV40796

AAV40796

AAV40796

AAV40796

AAV40796

AAV40796

(BAVU) BAYLOR COLLEGE MEDICINE.
 (GENE-) GENEMEDICINE INC.

Coleman M, Denayo FJ, Schwartz R;

WPI; 1998-333339/29.

New vector for expression of insulin-like growth factor-I -
 containing a skeletal alpha-actin gene promoter, IGF-I coding
 sequences and a 3' region from growth hormone 3'-UTR

Claim 16; Page 37-41; 115pp; English.

This sequence is the actual sequence for pIG0552 which is an example of
 a vector of the invention. The vector is for expression of a nucleic acid
 sequence in a cell, and comprises: (a) a nucleic acid cassette containing
 a sequence encoding insulin-like growth factor-I (IGF-I); (b) a
 5' flanking region including one or more sequences necessary for a
 expression of the nucleic acid cassette, including a promoter from a
 skeletal alpha-actin gene; (c) a linker connecting the 5' flanking region
 to a nucleic acid, the linker having a position for inserting the nucleic
 acid cassette, and lacking the coding sequence of a gene with which it is
 naturally associated; and (d) a 3' flanking region, including a
 3' untranslated region or a 3' non coding region or both, where the
 3' flanking region is 3' to the position for inserting the nucleic acid
 cassette and comprises a sequence from a growth hormone 3'-UTR. The
 vector can provide for efficient IGF-I expression, particularly in gene
 therapy. It can be used for the delivery of IGF-I for treating diseases
 such as muscle atrophy, diabetes, neuropathy, osteoporosis, and growth
 disorders. They can be used for treating peripheral neuropathies and
 resulting from diabetes, genetic disease such as Type I or Type II
 diabetes, genetic disease such as Chacot-marie-tooth disease, AIDS,
 atherosclerosis, atherosclerotic, cardiovascular, cerebrovascular, or
 peripheral vascular disease, haemophilia, inflammation and side-effects
 from anti-cancer and anti-viral drugs. The vectors can also be used to
 create transgenic animals for research or livestock improvement.

Sequence 3599 BP; 733 A; 1098 C; 1042 G; 726 T; 0 other;

Query Match 30.4%; Score 1603.8; DB 19; Length 3599;

Best Local Similarity 82.8%; Pred. No. 1.1e-239; Indels 392; Gaps 1;

Matches 2022; Conservative 0; Mismatches 27;

QY 2706 CTAGAGCTAATCATGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2765
 DB 56 CTTGGCGTAAATCATGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 115
 QY 2766 ACACAAATACGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2825
 DB 116 ACACAAATACGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 175
 QY 2826 ACTCAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2885
 DB 176 ACTCAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 235
 QY 2886 GCTGCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2945
 DB 236 GCTGCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 295
 QY 2946 GCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3005
 DB 296 GCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 355
 QY 3006 TCATGAGGCGGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3065
 DB 356 TCATGAGGCGGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 415
 QY 3066 GTGAGCAAAAAGCCAGCAAAAAGCCAGCAAAAAGCCAGCAAAAAGCCAGCAAAAAGCCAG 3125
 DB 416 GTGAGCAAAAAGCCAGCAAAAAGCCAGCAAAAAGCCAGCAAAAAGCCAGCAAAAAGCCAG 475
 QY 3126 CCATAGGCTCCG 3185

Actual sequence of IGF-I containing plasmid pIG0552.

IGF-I; insulin-like growth factor I; skeletal alpha-actin gene promoter;
 muscle atrophy; diabetes; osteoporosis; growth disorder; therapy; AIDS;
 Chacot-marie-tooth disease; atherosclerosis; haemophilia; neuropathy; ss.

Synthetic.

W09824922-A1.

11-JUN-1998.

01-DEC-1997; 97WO-0521852.

19-NOV-1997; 97US-0974572.

02-DEC-1996; 96US-0031539.

Db 476 CCGTATGCTCCGCCCCCTTACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCG 535
Qy 3186 AAACCCGACAGGACTATAAGATACACAGCGCTTCCCGCTGGAAGCTCCCTCGTGGGCTC 3245
Db 536 AAACCCGACAGGACTATAAGATACACAGCGCTTCCCGCTGGAAGCTCCCTCGTGGGCTC 595
Qy 3246 TCCGTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGCGGAAGCGT 3305
Db 596 TCCGTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGCGGAAGCGT 655
Qy 3306 GGGCTTTCTCATAGCTCAGCTGTAGTATCTCAGTTCGGGTAGGTGCTGCTCCAA 3365
Db 656 GGGCTTTCTCATAGCTCAGCTGTAGTATCTCAGTTCGGGTAGGTGCTGCTCCAA 715
Qy 3366 GCTGGGCTGTGTACGAACCCCGCTTACGCGGATACCTGTCGCGCTTATCCCGTAACTA 3425
Db 716 GCTGGGCTGTGTACGAACCCCGCTTACGCGGATACCTGTCGCGCTTATCCCGTAACTA 775
Qy 3426 TCGTCTTACGCTCAACCCGCTAAGACACAGCTTATCGCCACTGCGACAGCCCACTGGTAA 3485
Db 776 TCGTCTTACGCTCAACCCGCTAAGACACAGCTTATCGCCACTGCGACAGCCCACTGGTAA 835
Qy 3486 CAGATTAGCAGAGCAGGATGTAGGCGGTGCTACAGATCTTGAAGTGGTGGCTAA 3545
Db 836 CAGATTAGCAGAGCAGGATGTAGGCGGTGCTACAGATCTTGAAGTGGTGGCTAA 895
Qy 3546 CTACGGCTACACTAGAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGGTACCTT 3605
Db 896 CTACGGCTACACTAGAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGGTACCTT 955
Qy 3606 CGGAAAGAGTGGTACGCTTGTATCCGGCAACACACACCGCTGCTGAGCGGTGTTT 3665
Db 956 CGGAAAGAGTGGTACGCTTGTATCCGGCAACACACACCGCTGCTGAGCGGTGTTT 1015
Qy 3666 TTTTGTGTAAGCAGCAGATTAACGCGCAGAAAGAGTCTCAAGAGATCCTTTGAT 3725
Db 1016 TTTTGTGTAAGCAGCAGATTAACGCGCAGAAAGAGTCTCAAGAGATCCTTTGAT 1075
Qy 3726 CTTTCTACGGGTCTGACGCTCAGTGGAAAGAAACCTCAGTTAAGGGATTTTGTCTAT 3785
Db 1076 CTTTCTACGGGTCTGACGCTCAGTGGAAAGAAACCTCAGTTAAGGGATTTTGTCTAT 1093
Qy 3786 GAGATTATCTGACCAAGAGCGCCATCTGCGCTCCCACTCTCTGAGTTCGGGGGCAATG 3845
Db 1094 ----- 1093
Qy 3846 GATGCGCGATAGCGCTGCTGCTTTCTGATGCGGAGGATTTGCACTGCGGCTAGAA 3905
Db 1094 ----- 1093
Qy 3906 CTCGCGAGGTCTCCAGCTCAGGACAGCTGAACCACTCGGAGGGGATCGAGCCC 3965
Db 1094 ----- 1093
Qy 3966 GGGGTGGGGAAGAACTCCAGCATGAGATCCCGCGCTGAGGATCATCAGCGCGGCTC 4025
Db 1094 ----- 1093
Qy 4026 CCGGAAACGATTCCGAAGCCCAACCTTTATAGAGGCGCGGTGGAAATCGAAATCTCG 4085
Db 1094 ----- 1093
Qy 4086 TGATGGCAGGTGGGCGTCTGCTTGTGCTCGTCAATTTCCGAACCCAGATCCCGCTCAGAAG 4145
Db 1094 ----- CGCTCAGAAG 1103
Qy 4146 AACTCGTCAAGAGGCGATAGAGCGATGCGCTCGAATCGGAGCGGCGATACCGTAA 4205
Db 1104 AACTCGTCAAGAGGCGATAGAGCGATGCGCTCGAATCGGAGCGGCGATACCGTAA 1163
Qy 4206 AGCAGAGGAGCGGTACGCCATTCGCGGCTCTCAGCAATATCACGGGTAGCC 4265

Db 1164 AGCAGAGGAGCGGTACGCCCATTTGCGCCGCAAGCTCTTCAGCAATATCACGGGTAGCC 1223
Qy 4266 AAGCTATATCTCTGTAGCGGTTCGGCCACACCCAGCGCGGCACAGTGCATGAATCAGAA 4325
Db 1224 AAGCTATATCTCTGTAGCGGTTCGGCCACACCCAGCGCGGCACAGTGCATGAATCAGAA 1283
Qy 4326 AAGCGGCGATTTTCCACCATGATATTGGCAAGCAGGATCGCCATCGGTCAAGAGAGA 4385
Db 1284 AAGCGGCGATTTTCCACCATGATATTGGCAAGCAGGATCGCCATCGGTCAAGAGAGA 1343
Qy 4386 TCCGCGCGTGGGATACGGCGCTTGAAGCTTGGCAACAGTTCGCTGCGGCGGAGCCCC 4445
Db 1344 TCCGCGCGTGGGATACGGCGCTTGAAGCTTGGCAACAGTTCGCTGCGGCGGAGCCCC 1403
Qy 4446 TGATGCTCTTCTGTCAGATCATCTGATCGAACAAGCGGCTTCCATCCGAGTACGTCT 4505
Db 1404 TGATGCTCTTCTGTCAGATCATCTGATCGAACAAGCGGCTTCCATCCGAGTACGTCT 1463
Qy 4506 CGCTCGATCGGATGTTTCGCTTGGTTCGAATCGGAGGTAGCGGATCAAGCGTATGC 4565
Db 1464 CGCTCGATCGGATGTTTCGCTTGGTTCGAATCGGAGGTAGCGGATCAAGCGTATGC 1523
Qy 4566 AGCGCGCGATTTGCATCAGCCATGATGATATCTTCTCGGAGGAGCAAGGTGAGATGAC 4625
Db 1524 AGCGCGCGATTTGCATCAGCCATGATGATATCTTCTCGGAGGAGCAAGGTGAGATGAC 1583
Qy 4626 AGGAGATCTCTGCGCGGCACTTCCGCCAATAGCAGCGGCTTCCCGCTTCAGTGACA 4685
Db 1584 AGGAGATCTCTGCGCGGCACTTCCGCCAATAGCAGCGGCTTCCCGCTTCAGTGACA 1643
Qy 4686 ACCTCGACACAGCTTCGCGGAGGAGGAGCGGCTGCGGCGGAGGAGGAGGAGGAGGAGG 4745
Db 1644 ACCTCGACACAGCTTCGCGGAGGAGGAGGAGCGGCTGCGGCGGAGGAGGAGGAGGAGG 1703
Qy 4746 TCGTCTGTCAGTTCAATCAGGCGACCGGACAGCTGCTTTGAACAAAGAAACCGGCGCG 4805
Db 1704 TCGTCTGTCAGTTCAATCAGGCGACCGGACAGCTGCTTTGAACAAAGAAACCGGCGCG 1763
Qy 4806 CCGTGGCTGACAGCGGAAACACGCGGCGATCAGCAGCGGATTTGCTGTTGCGCCAG 4865
Db 1764 CCGTGGCTGACAGCGGAAACACGCGGCGATCAGCAGCGGATTTGCTGTTGCGCCAG 1823
Qy 4866 TATAGCCGATAGCTCTCCACCGAGCGGCGGAGAACTGGTGCAATCCATCTTGT 4925
Db 1824 TATAGCCGATAGCTCTCCACCGAGCGGCGGAGAACTGGTGCAATCCATCTTGT 1883
Qy 4926 TCAATCATCGGAAACGATCTCTCTCTTGTATCAGATCTTGAATCCCTGCGCCAT 4985
Db 1884 TCAATCATCGGAAACGATCTCTCTCTTGTATCAGATCTTGAATCCCTGCGCCAT 1943
Qy 4986 CAGATCTTGGCGGCAAGAAAGCCATCCAGTTTACTTTGAGGCTTCCCAACCTTACCA 5045
Db 1944 CAGATCTTGGCGGCAAGAAAGCCATCCAGTTTACTTTGAGGCTTCCCAACCTTACCA 2003
Qy 5046 GAGGCGCGCGGAGTGGCAATTCGGTTCGCTGCTTCCATAAACCGGCGGAGTTCAGC 5105
Db 2004 GAGGCGCGCGGAGTGGCAATTCGGTTCGCTGCTTCCATAAACCGGCGGAGTTCAGC 2063
Qy 5106 TATCGGCATGTAAGCCCACTGCAAGTACCTGCTTCTCTT 5146
Db 2064 AACTGTTGGGAGGCGATCGGTGGGCGCTCTTCGCTATT 2104

RESULT 12
AAV50427
ID AAV50427 standard; DNA; 3600 BP.
XX
AC AAV50427;
XX
DT 26-OCT-1998 (first entry)
XX
DB Plasmid pIG0552 upper expected sequence nucleotides 1-3600.
XX

KW Human; IGF-1; insulin-like growth factor 1; urinary incontinence;
 KW gene therapy; neurotrophic factor; ss.
 XX Synthetic.
 OS Homo sapiens.
 XX WO9833529-A1.
 XX 06-AUG-1998.
 XX 04-FEB-1998; 98WO-US02051.
 XX 04-FEB-1997; 97US-0036862.
 XX (GENE-) GENEMEDICINE INC.
 XX Coleman M;
 XX WPI; 1998-437184/37.
 XX Treatment of urinary incontinence - by delivering nucleic acid
 XX vector for expression of growth factor or neurotrophic factor in
 XX tissue(s)
 XX Disclosure; Page 103-105; 117pp; English.
 XX A method has been developed of treating urinary incontinence (UI) in
 XX mammals. The method comprises delivering a nucleic acid vector for the
 XX expression of a growth factor or neurotrophic factor in a tissue or
 XX tissues. The present sequence represents the expected upper sequence
 XX of plasmid pIG0552 nucleotides 1-3600 from the present invention. Due
 XX to the growth and stimulatory effects of growth factors and neurotrophic
 XX factors, introducing these factors to degenerated muscles in the
 XX urinary system can improve UI by enhancing both their integrity and
 XX neural innervation.
 XX Sequence 3600 BP; 733 A; 1099 C; 1042 G; 726 T; 0 other;
 XX
 XX Query Match 30.3%; Score 1602.2; DB 19; Length 3600;
 XX Best Local Similarity 82.8%; Pred. No. 1.9e-239;
 XX Matches 2021; Conservative 0; Mismatches 28; Indels 392; Gaps 1;
 XX
 XX 2706 CTAGAGCTAATGATGCTAGCTGTTCTCTGTTGTAATGTTATCCCGTCAATTC 2765
 XX 57 CTGGCGTATCATGTCATAGCTGTTCTCTGTTGTAATGTTATCCCGTCAATTC 116
 XX 2766 ACAACATACAGCGCGAGCATAAAGTGAAGCTGGGGTGCCTTAATGATGAGCTA 2825
 XX 117 ACAACATACAGCGCGAGCATAAAGTGAAGCTGGGGTGCCTTAATGATGAGCTA 176
 XX 2826 ACTCACATTAATTGGCTGCTCACTCCCGCTTTCCAGTCGGGAAACCTGTCGCCA 2885
 XX 177 ACTCACATTAATTGGCTGCTCACTCCCGCTTTCCAGTCGGGAAACCTGTCGCCA 236
 XX 2886 GCTGCAATTAATGATGCTGCTCACTCCCGCTTTCCAGTCGGGAAACCTGTCGCCA 2945
 XX 237 GCTGCAATTAATGATGCTGCTCACTCCCGCTTTCCAGTCGGGAAACCTGTCGCCA 296
 XX 2946 CGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 3005
 XX 297 CGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 356
 XX 3006 TCACTCAAAAGCGGTAATAGCTGTTATCCAGAGATCAGGGGTAACCGCAGGAAGACAT 3065
 XX 357 TCACTCAAAAGCGGTAATAGCTGTTATCCAGAGATCAGGGGTAACCGCAGGAAGACAT 416
 XX 3066 GTAGCAAAAGCGCAGCAAAAGCGCAGGAAACCGTAAAGGCGCGCTGCTGCGCTTTT 3125
 XX 417 GTAGCAAAAGCGCAGCAAAAGCGCAGGAAACCGTAAAGGCGCGCTGCTGCGCTTTT 476
 XX 3126 CCAATAGCTCCGCGCGCTGACGAGCATCACAAAATCGAGCTCAAGTCAGAGGTGGCG 3185
 XX 477 CCAATAGCTCCGCGCGCTGACGAGCATCACAAAATCGAGCTCAAGTCAGAGGTGGCG 536

QY 3186 AAACCCGACAGGACTATATAAGATACAGAGGCTTTCCCTCTGGAAGCTCCCTCGTGGCTC 3245
 Db 537 AAACCCGACAGGACTATATAAGATACAGAGGCTTTCCCTCTGGAAGCTCCCTCGTGGCTC 596
 QY 3246 TCCTGTTCCGACCTCGCGCTTACCGATACCTGTCGGCTTTCTCCCTCGGGAAGCT 3305
 Db 597 TCCTGTTCCGACCTCGCGCTTACCGATACCTGTCGGCTTTCTCCCTCGGGAAGCT 656
 QY 3306 GCGGCTTCTCATAGCTACGCTGTAGGTATCTCAGTTGCGGTGATGCTGCTCCAA 3365
 Db 657 GCGGCTTCTCATAGCTACGCTGTAGGTATCTCAGTTGCGGTGATGCTGCTCCAA 716
 QY 3366 GCTGGGCTGTGTGACAGCAACCCCGCTTACGCGCGAGCGCTGCGCTTATCGGTAAC 3425
 Db 717 GCTGGGCTGTGTGACAGCAACCCCGCTTACGCGCGAGCGCTGCGCTTATCGGTAAC 776
 QY 3426 TCGTCTTCAGTCCAAACCGGTAAAGACAGCTTATCCCACTGGCAGCAGCAGCTGTA 3485
 Db 777 TCGTCTTCAGTCCAAACCGGTAAAGACAGCTTATCCCACTGGCAGCAGCAGCTGTA 836
 QY 3486 CAGGATTAGCAGCAGGATGATGAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAA 3545
 Db 837 CAGGATTAGCAGCAGGATGATGAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAA 896
 QY 3546 CTACGCTACACTAGAGAAAGTATTTGGTATCTGCGCTCTCTGCTGAAGCAGTTACCT 3605
 Db 897 CTACGCTACACTAGAGAAAGTATTTGGTATCTGCGCTCTCTGCTGAAGCAGTTACCT 956
 QY 3606 CGGAAAAGGTTGGTACCTTTCATCCGCAACAAACACCGCTGCTGAGCGGTGTT 3665
 Db 957 CGGAAAAGGTTGGTACCTTTCATCCGCAACAAACACCGCTGCTGAGCGGTGTT 1016
 QY 3666 TTTGTTTGAAGCAGCAGATTAACCGCAGAAAAAAGGATCTCAAGAGATCCCTTTGAT 3725
 Db 1017 TTTGTTTGAAGCAGCAGATTAACCGCAGAAAAAAGGATCTCAAGAGATCCCTTTGAT 1076
 QY 3726 CTTTCTACGGGCTCTGACGCTCAGTGAACGAAACTCAGTTAAGGATTTTGGTCAT 3785
 Db 1077 CTTTCTACGGGCTCTG----- 1094
 QY 3786 GAGATTTCGTGACCAAAAGCGGCATCTGCTCCCTCTGCTGAGTTCCGGGGCATG 3845
 Db 1095 ----- 1094
 QY 3846 GATCGCGGATAGCGCTGCTGTTTCTGGATGCCGACGGATTTGCACTGCCGGTAGAA 3905
 Db 1095 ----- 1094
 QY 3906 CTCGCGAGGTCGTCAGCCTCAGGACAGCTGAACCAACTCGCGAGGGGATCGAGCC 3965
 Db 1095 ----- 1094
 QY 3966 GGGGTGGCGGAAGAACTCCAGCATGAGATCCCGCGCTGGAGATCTCCAGCCGGCTC 4025
 Db 1095 ----- 1094
 QY 4026 CGGAAAACGATTCGAAGCCCAACCTTTTCATAGAAAGCGGGGTGGAATCGAAATCTCG 4085
 Db 1095 ----- 1094
 QY 4086 TGATGGCAGGTTGGCGCTCGCTTGGTGGTCAATTCGAACCCAGAGTCCCGCTCAGAG 4145
 Db 1095 -----CGCTCAGAG 1104
 QY 4146 AACTCGTCAAGAGCGGATAGAGCGGATGCGTGTGGAATCGGAGCGCGGATACCGTAA 4205
 Db 1105 AACTCGTCAAGAGCGGATAGAGCGGATGCGTGTGGAATCGGAGCGCGGATACCGTAA 1164
 QY 4206 AGCAGAGAGCGGATAGAGCGGATGCGTGTGGAATCGGAGCGCGGATACCGTAA 4265
 Db 1165 AGCAGAGAGCGGATAGAGCGGATGCGTGTGGAATCGGAGCGCGGATACCGTAA 1224

RESULT 14

AA88055
ID AAX88055 standard; DNA; 5707 BP.

XX AC AAX88055;

XX DT 08-SEP-1999 (first entry)

XX DE Plasmid pIG0335 DNA.

XX KW Plasmid pIG0335; expression vector; treatment; disease;

XX KW muscular disease; gene therapy; muscle atrophy; neurological;

XX KW clotting factor; atherosclerosis; arteriosclerotic; cardiovascular;

XX KW diabetes; transgenic animal; carcinogen; regulatory element;

XX KW livestock improvement; immune response; ds.

XX OS Synthetic.

XX FN US5925564-A.

XX PD 20-JUL-1999.

XX PF 07-JUN-1995; 95US-0472809.

XX PR 07-JUN-1995; 95US-0472809.

XX PR 06-NOV-1991; 91US-0789919.

XX PR 09-MAR-1994; 94US-0209846.

XX XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX PI DeMayo FU, O'Malley BW, Schwartz RJ;

XX DR WPI; 1999-418276/35.

XX PT New expression vector system useful for gene therapy

XX PS Disclosure; Fig 17-17F; 67pp; English.

XX CC This invention describes novel expression vector systems containing RNA
CC stability elements from 3' flanking sequences used for establishing
CC expression of a nucleic acid sequence within a tissue. The vectors also
CC facilitate enhanced expression in tissues and target expression with
CC tissue specificity. The expression vectors can be used to treat diseases
CC through gene therapy by targeting the vector to specific tissues.
CC Diseases that can be treated include muscle atrophy associated with
CC neurological, muscular or systemic disease, aging by causing tissues to
CC express trophic factors, haemophilia by causing tissues to express and
CC secrete clotting factors into the circulation, atherogenesis and
CC atherosclerotic cardiovascular, cerebrovascular or peripheral-vascular
CC disease by causing tissues to express factors involved in tissue
CC metabolism. They can be used to replace genes of inherited genetic
CC defects or acquired hormone deficiencies e.g. diabetes. To transform
CC cells to produce particular proteins or RNA in vitro. To create
CC transgenic animals which can be used for research into human diseases,
CC assessing novel therapeutic methods, assessing the effect of chemical
CC and physical carcinogens and for studying the effect of genes and genetic
CC regulatory elements or livestock improvement. They can be used to induce
CC an immune response. These vectors provide controlled expression of the
CC genes they carry and produce a significantly high level of expression.
CC Using 3'UTR sequences reduces the decay rates of the mRNAs encoded by
CC the vectors which causes increased expression.

XX SQ Sequence 5707 BP; 1221 A; 1634 C; 1576 G; 1276 T; 0 other;

Query Match 30.3%; Score 1602.2; DB 20; Length 5707;

Best Local Similarity 82.8%; Pred. No. 1.9e-239;

Matches 2021; Conservative 0; Mismatches 28; Indels 392; Gaps 1;

QY 2706 CTAGACGTATCATGTCATAGCTGTTTCTGTGTAATTTGTTATCCGCTCACAATTCC 2765

Db	3514	CTTGCGCTTAATCATGGTCATAGCTGTTTCTGTGTAATTTGTTATCCGCTCACAATTCC	3573
QY	2766	ACACAACATACGAGCGGAGCAATAAAGTAAAGCTTGGGTCCTTAATGAGTGAGCTA	2825
Db	3574	ACACAACATACGAGCGGAGCAATAAAGTAAAGCTTGGGTCCTTAATGAGTGAGCTA	3633
QY	2826	ACTCACATTAATTCGGTTGGCTCACTGCGCGCTTTCAGTCCGGAAACCTGTCGTCGA	2885
Db	3634	ACTCACATTAATTCGGTTGGCTCACTGCGCGCTTTCAGTCCGGAAACCTGTCGTCGA	3693
QY	2886	GCTCATTAATTAATTCGGTCAACCGCGGGGAGAGCGGTTTTCGTTATGGGCGCTCTTC	2945
Db	3694	GCTCATTAATTAATTCGGTCAACCGCGGGGAGAGCGGTTTTCGTTATGGGCGCTCTTC	3753
QY	2946	CGCTTCCTCGCTCACTGACGCTGCGCTCGCTCGCTTTCGGCTGCGCGGAGCGGTATCAGC	3005
Db	3754	CGCTTCCTCGCTCACTGACGCTGCGCTCGCTCGCTTTCGGCTGCGCGGAGCGGTATCAGC	3813
QY	3006	TCATCTCAAGGCGGTAAATACGTTATCCACAGATCAGGGGATACGAGAAAGACAT	3065
Db	3814	TCATCTCAAGGCGGTAAATACGTTATCCACAGATCAGGGGATACGAGAAAGACAT	3873
QY	3066	GTGAGCAAAAAGGCGCAGCAAAAAGGCGCAGCAACCTAAAAAGGCGGTTTTCGTCGGCTTTT	3125
Db	3874	GTGAGCAAAAAGGCGCAGCAAAAAGGCGCAGCAACCTAAAAAGGCGGTTTTCGTCGGCTTTT	3933
QY	3126	CCATAGGCTCGCGCCCTGACGAGCATCAAAAATCGAGCTCAAGTCAGAGGTGGCG	3185
Db	3934	CCATAGGCTCGCGCCCTGACGAGCATCAAAAATCGAGCTCAAGTCAGAGGTGGCG	3993
QY	3186	AAACCCGACAGGCTATAAAGATACGAGCGTTCCGCTGAGAGCTCCCTCGCGCTC	3245
Db	3994	AAACCCGACAGGCTATAAAGATACGAGCGTTCCGCTGAGAGCTCCCTCGCGCTC	4053
QY	3246	TCCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTTCCTTCGGGAAAGCT	3305
Db	4054	TCCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTTCCTTCGGGAAAGCT	4113
QY	3306	GGCGTTTCTCATAGCTCAGCTGTAAGTATCTCATGTTCCGTTAGGTGCTTCCTCCAA	3365
Db	4114	GGCGTTTCTCATAGCTCAGCTGTAAGTATCTCATGTTCCGTTAGGTGCTTCCTCCAA	4173
QY	3366	GCTGGCTGTGTGACGAAACCCCGCTTCAGCCGAGCGCTGCGCTTTCGCGTAATA	3425
Db	4174	GCTGGCTGTGTGACGAAACCCCGCTTCAGCCGAGCGCTGCGCTTTCGCGTAATA	4233
QY	3426	TCGTTTGAAGTCCAAACCCGCTTAAGACACGACTTATCGGCACTGGCAGCAGCACTGTAA	3485
Db	4234	TCGTTTGAAGTCCAAACCCGCTTAAGACACGACTTATCGGCACTGGCAGCAGCACTGTAA	4293
QY	3486	CAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAA	3545
Db	4294	CAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAA	4353
QY	3546	CTACGGCTACATAGAGACAGTATTTGGTATCTCGCTCTGCTGAGCCAGTACCTT	3605
Db	4354	CTACGGCTACATAGAGACAGTATTTGGTATCTCGCTCTGCTGAGCCAGTACCTT	4413
QY	3606	CGGAAAGAGTGGTGTAGCTTTGATCCGGCAAAACAAACCCGCTGGTAGCGGTGTT	3665
Db	4414	CGGAAAGAGTGGTGTAGCTTTGATCCGGCAAAACAAACCCGCTGGTAGCGGTGTT	4473
QY	3666	TTTTGTTTGAAGCAGCAGATTAACGCGCAGAAAAAAGGATCTCAAGAAGATCTTTGAT	3725
Db	4474	TTTTGTTTGAAGCAGCAGATTAACGCGCAGAAAAAAGGATCTCAAGAAGATCTTTGAT	4533
QY	3726	CTTTTCTACGGGCTCTGACGCTCAGTGGAAACGAAACTCAGCTTAAGGATTTTGGTCAT	3785
Db	4534	CTTTTCTACGGGCTCTGACGCTCAGTGGAAACGAAACTCAGCTTAAGGATTTTGGTCAT	4551
QY	3786	GAGATTATCGTCGACCAAGACCGGCTATCGTGCTCCCACTCTCGAGTTCGGGGGATG	3845

Db 120 ACACACATACGACCGGAGCATAAAGTGTAAAGCCTGGGTGCTTAATGAGTACGCTA 179
Qy 2826 ACTCACATTAATTTGGCTTGGCTCACTGCGCCGCTTTCCAGTGGGAAACCTGTGTGCTCA 2885
Db 180 ACTCACATTAATTTGGCTTGGCTCACTGCGCCGCTTTCCAGTGGGAAACCTGTGTGCTCA 239
Qy 2886 GCTGCAATTAATTAATGAAATCGGCAACGCGCGGGAGAGCGGTTTGGCTATTGGGCGCTTTC 2945
Db 240 GCTGCAATTAATTAATGAAATCGGCAACGCGCGGGAGAGCGGTTTGGCTATTGGGCGCTTTC 299
Qy 2946 CGCTTCTCGCTCACTGACTCGCTGCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3005
Db 300 CGCTTCTCGCTCACTGACTCGCTGCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 359
Qy 3006 TCACTCAAGCGCGTAAATCGGTTATCCACAGAAATCAGGGGATAACGAGGAAAGAAACAT 3065
Db 360 TCACTCAAGCGCGTAAATCGGTTATCCACAGAAATCAGGGGATAACGAGGAAAGAAACAT 419
Qy 3066 GTGAGCAAAAGCGCAGCAAAAGCGCAGCAAAAGCGTAAAGGCGCGGTTGCTGGGCTTTT 3125
Db 420 GTGAGCAAAAGCGCAGCAAAAGCGCAGCAAAAGCGTAAAGGCGCGGTTGCTGGGCTTTT 479
Qy 3126 CCAATAGCTCCGCGCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAAGAGTGGCG 3185
Db 480 CCAATAGCTCCGCGCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAAGAGTGGCG 539
Qy 3186 ABAACCGCAGAGTAAAGATACAGGCGTTTCCCTCGAAGCTCCCTGCTGCGCTC 3245
Db 540 ABAACCGCAGAGTAAAGATACAGGCGTTTCCCTCGAAGCTCCCTGCTGCGCTC 599
Qy 3246 TCTGTTCCGACCTCCGCTTACCGGATACGCTGCGCTTTCCTCCCTCGGGAAGCGT 3305
Db 600 TCTGTTCCGACCTCCGCTTACCGGATACGCTGCGCTTTCCTCCCTCGGGAAGCGT 659
Qy 3306 GCGCTTCTCATAGCTCAAGTACGCTGAGTATCTAGTTCGCTGAGTACGCTCCCTCAA 3365
Db 660 GCGCTTCTCATAGCTCAAGTACGCTGAGTATCTAGTTCGCTGAGTACGCTCCCTCAA 719
Qy 3366 GCTGGGCTGTGTACGAAACCCCGCTTACGCGGCTGCGCTTATCCGCTTAACTA 3425
Db 720 GCTGGGCTGTGTACGAAACCCCGCTTACGCGGCTGCGCTTATCCGCTTAACTA 779
Qy 3426 TCGTCTGAGTCAACCCGCTGAGCAGCTTATCGCCTGCGCTGAGTACGCTGCTGCTGCT 3485
Db 780 TCGTCTGAGTCAACCCGCTGAGCAGCTTATCGCCTGCGCTGAGTACGCTGCTGCTGCT 839
Qy 3486 CAGGATAGCAGAGGATGTAGCGGCTGCTACAGAGTTCCTTGAAGTGGTGGCTTAA 3545
Db 840 CAGGATAGCAGAGGATGTAGCGGCTGCTACAGAGTTCCTTGAAGTGGTGGCTTAA 899
Qy 3546 CTAGGCTACACTAGAGACAGTATTTGGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3605
Db 900 CTAGGCTACACTAGAGAGCAGTATTTGGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 959
Qy 3606 CGGAAAGAGTGGTAGCTTCTTGAATCGGCAAAACCAACCGCTGCTGCTGCTGCTGCTGCT 3665
Db 960 CGGAAAGAGTGGTAGCTTCTTGAATCGGCAAAACCAACCGCTGCTGCTGCTGCTGCTGCT 1019
Qy 3666 TTTTGTTCAGCAGCAGTACGCGCAGAAAGAGTCTCAAGAGATCTTCTGAT 3725
Db 1020 TTTTGTTCAGCAGCAGTACGCGCAGAAAGAGTCTCAAGAGATCTTCTGAT 1079
Qy 3726 CTTTCTACCGGCTGCTGAGCTCAGTGGAAACGAAACCTCAGTTAAGGATTTTGGTAT 3785
Db 1080 CTTTCTACCGGCTGCTGAGCTCAGTGGAAACGAAACCTCAGTTAAGGATTTTGGTAT 1097
Qy 3786 GAGATTATCTGACCAAAAGCGCCATCGTGCTCCCTCCCTGCTGCTGCTGCTGCTGCTGCT 3845
Db 1098 ----- 1097
Qy 3846 GATGCGCGGATAGCCGCTGCTGCTTCTGAGTACGAGGATTTGCACTGCGGCTAGAA 3905
Db 1098 ----- 1097

Qy 3906 CTTCCGCGAGGCTGCTCCAGCCTCAGGACGAGCTGAACCAACTCGGAGGGGATCGAGCCC 3965
Db 1098 ----- 1097
Qy 3966 GGGGTGGGCGAAGAACTCCAGCATGAGATCCCGGCTGGAGGATCATCCAGCGGGCTC 4025
Db 1098 ----- 1097
Qy 4026 CCGGAAACGATTCGGAAGCCCAACCTTTATAGAGGCGGCTGGATCGAATCTGATCTCG 4085
Db 1098 ----- 1097
Qy 4086 TGATGGCAGGTTGGGCGTGGCTTGGTGGTCAATTCGAACCCAGAGTCCCGCTCAGAAG 4145
Db 1098 -----CGTCAAGAG 1107
Qy 4146 AACTGCTCAGAGGCGATAGAGCGGATGGCTGCTGCAATCGGAGCGGGGATACCGTAA 4205
Db 1108 AACTGCTCAGAGGCGATAGAGCGGATGGCTGCTGCAATCGGAGCGGGGATACCGTAA 1167
Qy 4206 AGCAGAGGAGCGGTCAAGCCATTTGCGCGCAAGCTCTTCAGCAATATCAAGCGGTAGCC 4265
Db 1168 AGCAGAGGAGCGGTCAAGCCATTTGCGCGCAAGCTCTTCAGCAATATCAAGCGGTAGCC 1227
Qy 4266 AAGCGGCGCATTTTCCACCATGATATTCGGCAAGCAGGATCGCCATGGGTCAAGCAGGA 4325
Db 1228 AAGCGGCGCATTTTCCACCATGATATTCGGCAAGCAGGATCGCCATGGGTCAAGCAGGA 1287
Qy 4326 TCTGCTCTTCTGCTCAGATCATCTGATCGAACAAGCGGCTTCCATCCGAGTACGTGCT 4385
Db 1408 TCTGCTCTTCTGCTCAGATCATCTGATCGAACAAGCGGCTTCCATCCGAGTACGTGCT 1467
Qy 4386 TCTGCTCTTCTGCTCAGATCATCTGATCGAACAAGCGGCTTCCATCCGAGTACGTGCT 4445
Db 1408 TCTGCTCTTCTGCTCAGATCATCTGATCGAACAAGCGGCTTCCATCCGAGTACGTGCT 1467
Qy 4446 TGATGCTCTTCTGCTCAGATCATCTGATCGAACAAGCGGCTTCCATCCGAGTACGTGCT 4505
Db 1408 TGATGCTCTTCTGCTCAGATCATCTGATCGAACAAGCGGCTTCCATCCGAGTACGTGCT 1467
Qy 4506 CGCTCGATCGGATGTTTCCGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4565
Db 1468 CGCTCGATCGGATGTTTCCGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1527
Qy 4566 AGCCCGCGCATTCGATCGCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4625
Db 1528 AGCCCGCGCATTCGATCGCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1587
Qy 4626 AGGAGATCTGCGCGGCGCATTCGCGCAATAGCAGGATCGGCTTCCGCTTCCGCTTCCGCT 4685
Db 1588 AGGAGATCTGCGCGGCGCATTCGCGCAATAGCAGGATCGGCTTCCGCTTCCGCTTCCGCT 1647
Qy 4686 AGCTCGAGCAGGTGCGCAAGGAAACCGCTGCTGCGCAGCAGCAGTACGCGCTGCTGCT 4745
Db 1648 AGCTCGAGCAGGTGCGCAAGGAAACCGCTGCTGCGCAGCAGCAGTACGCGCTGCTGCT 1707
Qy 4746 TCGTCTCTGAGTTCATTCAGGCGCACCGGACAGGCTGCTTGAACAAAGAAACCGGCGCG 4805
Db 1708 TCGTCTCTGAGTTCATTCAGGCGCACCGGACAGGCTGCTTGAACAAAGAAACCGGCGCG 1767
Qy 4806 CCGTCTGCTGAGCAGCGGAAACCGCGGCGATCAGAGCAGCGGATGCTGCTGCTGCTGCTGCT 4865
Db 1768 CCGTCTGCTGAGCAGCGGAAACCGCGGCGATCAGAGCAGCGGATGCTGCTGCTGCTGCTGCT 1827
Qy 4866 TCATAGCGCAATAGGCTCTCCACCCAGCGGCGGAGAACTGCTGCTGCTGCTGCTGCTGCT 4925
Db 1828 TCATAGCGCAATAGGCTCTCCACCCAGCGGCGGAGAACTGCTGCTGCTGCTGCTGCTGCT 1887
Qy 4926 TCATAGCGCAATAGGCTCTCCACCCAGCGGCGGAGAACTGCTGCTGCTGCTGCTGCTGCT 4985
Db 1888 TCATAGCGCAATAGGCTCTCCACCCAGCGGCGGAGAACTGCTGCTGCTGCTGCTGCTGCT 1947

Tue Feb 17 08:36:51 2004

us-09-921-143-36.rng

QY	4986	CAGATCCTTGGCGGCAAGAGCCATCCAGTTTACTTTGACGGCTTCCCAACCTTACCA	5045
Db	1948	CAGATCCTTGGCGGCAAGAGCCATCCAGTTTACTTTGACGGCTTCCCAACCTTACCA	2007
QY	5046	GAGGGCGCCCCAGCTGGCAATTCGGTTGCTGCTCCATAAAACGGCCAGTCTAGC	5105
Db	2008	GAGGGCGCCCCAGCTGGCAATTCGGTTGCTGCTCCATAAAACGGCCAGTCTAGC	2067
QY	5106	TA	5107
Db	2068	AA	2069

Search completed: February 15, 2004, 11:11:36
Job time : 1314 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 01:53:42 ; Search time 272 Seconds
(without alignments)
8572.894 Million cell updates/sec

Title: US-09-921-143-36

Perfect score: 5283
Sequence: 1 aagttgacottatcgact.....tgagtgcttcgcgcagcgtg 5283

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 563978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B COMB.seq: *
3: /cgn2_6/ptodata/2/ina/5A COMB.seq: *
4: /cgn2_6/ptodata/2/ina/5B COMB.seq: *
5: /cgn2_6/ptodata/2/ina/FACTUS COMB.seq: *
6: /cgn2_6/ptodata/2/ina/FACTUS1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2240	42.4	8349	4	US-09-186-002-16
2	1793.4	33.9	4800	4	US-09-554-928-1
3	1652.4	31.3	5594	4	US-09-380-190A-29
4	1618.2	30.6	6561	4	US-09-380-190A-30
5	1602.2	30.3	5707	2	US-08-472-809B-8
6	1571.6	29.7	4665	3	US-08-948-378A-7
7	1571.6	29.7	4665	3	US-09-169-425C-7
8	1571.6	29.7	4665	4	US-09-759-960-7
9	1570	29.7	4518	4	US-09-380-190A-26
10	1570	29.7	4896	4	US-09-533-220A-4
11	1570	29.7	6139	2	US-08-751-767A-7
12	1569.8	29.7	8797	2	US-08-723-306-6
13	1569.8	29.7	8797	5	PCT-US96-10041-6
14	1569.8	29.7	11093	2	US-08-723-306-5
15	1569.8	29.7	11093	5	PCT-US96-10041-5
16	1561.4	28.4	6795	4	US-09-380-190A-22
17	1501.4	28.4	8574	4	US-09-554-572-3
18	1501.4	28.4	9093	4	US-09-380-190A-23
19	1501.4	28.4	9093	4	US-09-554-572-2
20	1501.4	28.4	9145	4	US-09-554-572-1
21	1501	28.4	4958	4	US-09-380-190A-20
22	1501	28.4	5724	4	US-09-380-190A-18
23	1501	28.4	5724	4	US-09-380-190A-19
24	1501	28.4	6225	4	US-09-380-190A-1
25	1501	28.4	6321	4	US-09-380-190A-17
26	1501	28.4	6359	4	US-09-380-190A-15
27	1501	28.4	6891	4	US-09-380-190A-16

C	28	1501	28.4	8657	4	US-09-380-190A-14	Sequence 14, Appl
	29	1490.4	28.2	3974	3	US-09-026-343-33	Sequence 33, Appl
	30	1490.4	28.2	3974	3	US-09-042-105-16	Sequence 16, Appl
	31	1490.4	28.2	3974	3	US-09-044-856A-7	Sequence 7, Appl
	32	1490.4	28.2	3974	3	US-09-023-082A-147	Sequence 147, Appl
	33	1490.4	28.2	3974	3	US-09-044-855A-7	Sequence 7, Appl
	34	1490.4	28.2	3974	3	US-09-078-670-4	Sequence 4, Appl
	35	1490.4	28.2	3974	4	US-09-026-408-14	Sequence 14, Appl
	36	1490.4	28.2	3974	4	US-09-362-871-33	Sequence 33, Appl
	37	1490.4	28.2	3974	4	US-09-627-154-4	Sequence 4, Appl
	38	1490.4	28.2	3974	4	US-09-027-287-50	Sequence 50, Appl
	39	1490.4	28.2	3974	4	US-09-437-602-4	Sequence 4, Appl
	40	1490.4	28.2	3974	4	US-09-252-656B-50	Sequence 50, Appl
	41	1490.4	28.2	3984	3	US-09-044-795A-10	Sequence 10, Appl
	42	1490.4	28.2	3984	4	US-09-725-460A-10	Sequence 10, Appl
	43	1490.4	28.2	4208	3	US-09-044-796A-9	Sequence 9, Appl
	44	1490.4	28.2	4208	4	US-09-725-460A-9	Sequence 9, Appl
	45	1490.4	28.2	4208	4	US-09-689-693-37	Sequence 37, Appl

ALIGNMENTS

RESULT 1

US-09-186-002-16
; Sequence 16, Application US/09186002B
; Patent No. 6489542
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Romano, Charles P.
; TITLE OF INVENTION: Improved Method for Transforming Plants to Express
; TITLE OF INVENTION: delta-Endotoxins
; FILE REFERENCE: 38-21(13547) US Pat No. 6489542 09/186,002
; CURRENT APPLICATION NUMBER: US/09/186,002B
; CURRENT FILING DATE: 1998-11-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 16
; LENGTH: 8349
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: gene
; LOCATION: (3666)..(5573)
; OTHER INFORMATION: completely synthesized
US-09-186-002-16

Query Match	42.4%	Score 2240;	DB 4;	Length 8349;
Best Local Similarity	93.4%	Pred. No. 0;		
Matches 2409;	Conservative 0;	Mismatches 25;	Indels 144;	Gaps 1;
Qy	2706	CTAGACGTAAATCATGGTCATAGCTGTTTCTCTGTGGAATTTGTTATCCGTCACATTC	2765	
Db	5911	CTTGGCGTAATCATGGTCATAGCTGTTTCTCTGTGGAATTTGTTATCCGTCACATTC	5970	
Qy	2766	ACCAACATAGGACCGGAGAGCTAAAGCTGGGTCCTTAATGAGTGAGCTA	2825	
Db	5971	ACCAACATAGGACCGGAGAGCTAAAGCTGGGTCCTTAATGAGTGAGCTA	6030	
Qy	2826	ACTCACATTAATTCGTTGGCTCACTGCCCGCTTTTCCAGTCGGGAAACCTGTCGTGCCA	2885	
Db	6031	ACTCACATTAATTCGTTGGCTCACTGCCCGCTTTTCCAGTCGGGAAACCTGTCGTGCCA	6090	
Qy	2886	GCTGCATTAATGATCGGCCAAACCGCGGGAGAGCGGTTTCGTTATTCGGGCTCTTC	2945	
Db	6091	GCTGCATTAATGATCGGCCAAACCGCGGGAGAGCGGTTTCGTTATTCGGGCTCTTC	6150	
Qy	2946	CGCTTCCTCGCTCACTCACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT	3005	
Db	6151	CGCTTCCTCGCTCACTCACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT	6210	
Qy	3006	TCATCTAAGCGCGGTAATACGGTTATCCAGATCAGGAGGATAACGACGAGAACAT	3065	

Db 6211 TCACTCAAAGCGCGTAAATACGGTTATCCACAGAAATCAGGGGATACCGCAGGAAAGACAT 6270
Qy 3066 GTGAGCAAAAGGCGCAGCAAGGCCAGGAAACCGTAAAGAGCGCGTCTGCTGCGCTTTT 3125
Db 6271 GTGAGCAAAAGGCGCAGCAAAAGGCCAGGAAACCGTAAAGAGCGCGTCTGCTGCGCTTTT 6330
Qy 3126 CCATAGGCTCCGCCCCCTGACAGAGCATCAAAAAATCGACGCTCAAGTCAGAGTGCGG 3185
Db 6331 CCATAGGCTCCGCCCCCTGACAGAGCATCAAAAAATCGACGCTCAAGTCAGAGTGCGG 6390
Qy 3186 ARAACCGACAGGACTATAAGATACCAAGCGGTTTCCCTGGAAGCTCCCTGCTGCGCTC 3245
Db 6391 ARAACCGACAGGACTATAAGATACCAAGCGGTTTCCCTGGAAGCTCCCTGCTGCGCTC 6450
Qy 3246 TCCTGTTCCGACCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGCGAAGCGT 3305
Db 6451 TCCTGTTCCGACCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGCGAAGCGT 6510
Qy 3306 GCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAA 3365
Db 6511 GCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAA 6570
Qy 3366 GCTGGGCTGTGTGACGAAACCCCGCTTACGCCGACCGCTGCGCTTATCCGCTAACTA 3425
Db 6571 GCTGGGCTGTGTGACGAAACCCCGCTTACGCCGACCGCTGCGCTTATCCGCTAACTA 6630
Qy 3426 TCGTCTTAGTCCAAACCGGTAGACACGACTTATCGGCACCTGGCAGCAGCAGCTGTTAA 3485
Db 6631 TCGTCTTAGTCCAAACCGGTAGACACGACTTATCGGCACCTGGCAGCAGCAGCTGTTAA 6690
Qy 3486 CAGATTAGCAGACGAGGTATGTAGGCGGTGTACAGAGTTCTTCAAGTGGTGGCTTAA 3545
Db 6691 CAGATTAGCAGACGAGGTATGTAGGCGGTGTACAGAGTTCTTCAAGTGGTGGCTTAA 6750
Qy 3546 CTACGGCTACACTAGAAAGACAGTATTTGGTATCTGCGCTCTGCTCAAGCCAGTTACCTT 3605
Db 6751 CTACGGCTACACTAGAAAGACAGTATTTGGTATCTGCGCTCTGCTCAAGCCAGTTACCTT 6810
Qy 3606 CGGAAAGAGTTGGTAGCTTGTATCGGCAACCAACCAACCAACCAACCAACCAACCAAC 3665
Db 6811 CGGAAAGAGTTGGTAGCTTGTATCGGCAACCAACCAACCAACCAACCAACCAACCA 6870
Qy 3666 TTTTGTTCAGCAGCAGATTACCGCCAGAAAGAAAGATCTCAAGAGATCCTTTGAT 3725
Db 6871 TTTTGTTCAGCAGCAGATTACCGCCAGAAAGAAAGATCTCAAGAGATCCTTTGAT 6930
Qy 3726 CTCTTCTACGGGCTCAGCTCAGTGGACGAAACTCAGTTAAGGATTTTGGTCAAT 3785
Db 6931 CTCTTCTACGGGCTCAGCTCAGTGGACGAAACTCAGTTAAGGATTTTGGTCAAT 6990
Qy 3786 GAGATTATCGTCGACCAAGCGGCATCGTGCCTCCCTCCCTCCTGCAAGTTTCGGGGCATG 3845
Db 6991 GAGATTATCAAAAAGGATCTTCACTAGATCCTT----- 7024
Qy 3846 GATGGCGGATAGCGGCTGCTGGTTTCTGATGCCACGGATTGCACTGCCGCTAGAA 3905
Db 7025 ----- 7024
Qy 3906 CTCGCGAGGTCTGTCAGCCTCAGGACAGCTGAAACCACTCCGCGGGGATCGAGCCC 3965
Db 7025 -----TT 7026
Qy 3966 GGGGTGGCGAAGAACTCCAGATGAGATCCCGGCTGGAGGATCATCCAGCCGGGCTC 4025
Db 7027 GGGGTGGCGAAGAACTCCAGATGAGATCCCGGCTGGAGGATCATCCAGCCGGGCTC 7086
Qy 4026 CCGGAAACGATTCCGAGGCCAACCTTTATAGAGCGCGGCTGGAATCGAAATCTCG 4085
Db 7087 CCGGAAACGATTCCGAGGCCAACCTTTATAGAGCGCGGCTGGAATCGAAATCTCG 7146
Qy 4086 TGATGGCAGTTGGCGCTGCTGGTCAATTCGAAACCCCAAGATCCGCTCAGAAG 4145
Db 7147 TGATGGCAGTTGGCGCTGCTGGTCAATTCGAAACCCCAAGATCCGCTCAGAAG 7206

Qy 4146 AACTCGTCAAGAAAGCGGATAGAAAGCGGATCGCTGCGAATCGGAGCGGCGATACCGTAA 4205
Db 7207 AACTCGTCAAGAAAGCGGATAGAAAGCGGATCGCTGCGAATCGGAGCGGCGATACCGTAA 7266
Qy 4206 AGCAGAGGAGCGGTCAGCCCATTCGCGCCAGCTCTTTCAGCAATATCAGCGGTAGCC 4265
Db 7267 AGCAGAGGAGCGGTCAGCCCATTCGCGCCAGCTCTTTCAGCAATATCAGCGGTAGCC 7326
Qy 4266 AACGCTATGCTCTGATAGCGGTCGCGCACACCCAGCGCGGCACAGTCGATGAATCCAGAA 4325
Db 7327 AACGCTATGCTCTGATAGCGGTCGCGCACACCCAGCGCGGCACAGTCGATGAATCCAGAA 7386
Qy 4326 AAGCGGCATTTTCCACATGATATTCGGCAGCAGGATTCGCAATGGGTTCAGCAGCAGA 4385
Db 7387 AAGCGGCATTTTCCACATGATATTCGGCAGCAGGATTCGCAATGGGTTCAGCAGCAGA 7446
Qy 4386 TCCTTCGCGCTCGGCGATCGCGCTTCGAGCTCGGCAACAGTTTCGGTGGCGGAGCCCC 4445
Db 7447 TCCTTCGCGCTCGGCGATCGCGCTTCGAGCTCGGCAACAGTTTCGGTGGCGGAGCCCC 7506
Qy 4446 TGATGCTCTTCGTCAGATCATCTGATCGACAGACCGGCTTCATCCGAGTACGTGCT 4505
Db 7507 TGATGCTCTTCGTCAGATCATCTGATCGACAGACCGGCTTCATCCGAGTACGTGCT 7566
Qy 4506 CGCTCGATGCGATGTTTCGCTTGGTTCGAATCGGCGAGGTAGCCGATCAAGCGTATGC 4565
Db 7567 CGCTCGATGCGATGTTTCGCTTGGTTCGAATCGGCGAGGTAGCCGATCAAGCGTATGC 7626
Qy 4566 AGCGCGCATPTGATCAGCCATGATGATATCTTTCGCGCAGGAGCAGGTAGATGAC 4625
Db 7627 AGCGCGCATPTGATCAGCCATGATGATATCTTTCGCGCAGGAGCAGGTAGATGAC 7686
Qy 4626 AGGAGATCTCGCCCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCGCGCTTCACTGACA 4685
Db 7687 AGGAGATCTCGCCCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCGCGCTTCACTGACA 7746
Qy 4686 ACCTCGACACAGCTCGCAGAAAGCAACCGCTTCGCGCAGCAGTACGATAGCCGCTGCC 4745
Db 7747 ACCTCGACACAGCTCGCAGAAAGCAACCGCTTCGCGCAGCAGTACGATAGCCGCTGCC 7806
Qy 4746 TCGTCTCTGAGTTTCACTTCAGGGCACCGGACAGGTTCGCTTCGACAAAAGAACCGGGCGC 4805
Db 7807 TCGTCTCTGAGTTTCACTTCAGGGCACCGGACAGGTTCGCTTCGACAAAAGAACCGGGCGC 7866
Qy 4806 CCCTCGCTGACAGCGGCAACACCGCGCATCAGACAGCGCATTTGTGTGTGCCAG 4865
Db 7867 CCCTCGCTGACAGCGGCAACACCGCGCATCAGACAGCGCATTTGTGTGTGCCAG 7926
Qy 4866 TCATAGCGAATAGCTCTCCACCCAGCGCCCGAGNACCTGGTGCATTCATCTGT 4925
Db 7927 TCATAGCGAATAGCTCTCCACCCAGCGCCCGAGNACCTGGTGCATTCATCTGT 7986
Qy 4926 TCAATCATGCGAAGCATCTCATCTCTCTTTCGATCAGATCTTTCCTCCCTCGCCAT 4985
Db 7987 TCAATCATGCGAAGCATCTCATCTCTCTTTCGATCAGATCTTTCCTCCCTCGCCAT 8046
Qy 4986 CAGATCTTTCGCGGCAAGAAAGCATCCAGTTTACCTTCAGGGCTTCGACCTTACCA 5045
Db 8047 CAGATCTTTCGCGGCAAGAAAGCATCCAGTTTACCTTCAGGGCTTCGACCTTACCA 8106
Qy 5046 GAGGCGCCCGCAGCTGGCAATTCGCGTTCGCTTCGCTTCCATAAAACCGCCAGTCTAGC 5105
Db 8107 GAGGCGCCCGCAGCTGGCAATTCGCGTTCGCTTCGCTTCCATAAAACCGCCAGTCTAGC 8166
Qy 5106 TATCGCCATGTAAGCCCATCTGCAAGTACCTGCTTCTTTCGCTTGGCTTGGCTTCCCTT 5165
Db 8167 TATCGCCATGTAAGCCCATCTGCAAGTACCTGCTTCTTTCGCTTGGCTTGGCTTCCCTT 8226
Qy 5166 GTCCAGATAGCCAGTGTGATTCATTCGCGGTTCAGCACCGTTTCTGCGGACTGGCT 5225
Db 8227 GTCCAGATAGCCAGTGTGATTCATTCGCGGTTCAGCACCGTTTCTGCGGACTGGCT 8286

4867 CATAGCCGATAGCTCTCCACCCAGCGCCGAGAACCTGGTGCATCCATCTTGTT 4926
Db |||||
3236 CATAGCCGATAGCTCTCCACCCAGCGCCGAGAACCTGGTGCATCCATCTTGTT 3177
Qy |||||
4927 CAATCATCGGAAAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4986
Db |||||
3176 CAATCATCGGAAAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3117
Qy |||||
4987 AGATCTTGGGGGGAAGAAAGCCATCCAGTTTACTTTTTCAGGGCTTCCCAACCTTACCAG 5046
Db |||||
3116 AGATCTTGGGGGGAAGAAAGCCATCCAGTTTACTTTTTCAGGGCTTCCCAACCTTACCAG 3057
Qy |||||
5047 AGGCGGCCAGCTGCGCAATTCGGTTCGCTGTGTCGTGTCATATAAAACCGCCAGCTAGCT 5106
Db |||||
3056 AGGCGGCCAGCTGCGCAATTCGGTTCGCTGTGTCGTGTCATATAAAACCGCCAGCTAGCT 2997
Qy |||||
5107 ATCGCATGTAAGCCCACTGCAAGCTAGCTCTTCTCTTTGGCGCTTGGCTTCCCTTG 5166
Db |||||
2996 ATCGCATGTAAGCCCACTGCAAGCTAGCTCTTCTCTTTGGCGCTTGGCTTCCCTTG 2937
Qy |||||
5167 TCCAGATAGCCAGTAGCTGACATTCATCCGGGTGACGACCGCTTCTGCGGACTGGCTT 5226
Db |||||
2936 TCCAGATAGCCAGTAGCTGACATTCATCCGGGTGACGACCGCTTCTGCGGACTGGCTT 2877
Qy |||||
5227 TCTAGCTTCCGCTTCTTTAGCAGCCCTTGGCGCTGAGTCTTGGCGAGGGTG 5283
Db |||||
2876 TCTAGCTTCCGCTTCTTTAGCAGCCCTTGGCGCTGAGTCTTGGCGAGGGTG 2820

RESULT 3

US-09-380-190A-29/c
; Sequence 29, Application US/09380190A
; Patent No. 6410220

GENERAL INFORMATION:

APPLICANT: NATURE TECHNOLOGY CORPORATION, ET AL.
TITLE OF INVENTION: SELF-ASSEMBLING GENES, VECTORS AND USES
THEREOF

NUMBER OF SEQUENCES: 74

CORRESPONDENCE ADDRESS:

ADDRESSEE: MUEITING, RAASCH & GEBHARDT, P.A.

STREET: 119 NORTH FOURTH STREET, SUITE 203

CITY: MINNEAPOLIS

STATE: MINNESOTA

COUNTRY: USA

ZIP: 55401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/380,190A

FILING DATE: 26-Aug-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US98/03918

FILING DATE: 28-FEB-98

ATTORNEY/AGENT INFORMATION:

NAME: MUEITING, ANN M.

REGISTRATION NUMBER: 33,977

REFERENCE/DOCKET NUMBER: 228.00010201

TELECOMMUNICATION INFORMATION:

TELEPHONE: 612-305-1217

TELEFAX: 612-305-1228

INFORMATION FOR SEQ ID NO: 29:

SEQUENCE CHARACTERISTICS:

LENGTH: 594 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

SEQUENCE DESCRIPTION: SEQ ID NO: 29:

US-09-380-190A-29

Query Match 31.3%; Score 1652.4; DB 4; Length 5594;
Best Local Similarity 85.4%; Pred. No. 0;
Matches 1916; Conservative 0; Mismatches 261; Indels 67; Gaps 4;

Qy 3107 CGCGTTCGCTGGCGTTTTCATAGGCTCCGCCCTCGAGGACATCACAAAATCGAC 3166
Db 5591 CGCGTTCGCTGGCGTTTTCATAGGCTCCGCCCTCGAGGACATCACAAAATCGAC 5532
Qy 3167 GCTCAAGTCAGAGGTGGGAAACCCGACAGACTATAAAGATACAGCGCTTCCCGCTG 3226
Db 5531 GCTCAAGTCAGAGGTGGGAAACCCGACAGACTATAAAGATACAGCGCTTCCCGCTG 5472
Qy 3227 GAAGTCCCTCGTGGCTCTCTGTTCCGACCTTCGCGCTTACCGGATACGCTGCGGCT 3286
Db 5471 GAAGTCCCTCGTGGCTCTCTGTTCCGACCTTCGCGCTTACCGGATACGCTGCGGCT 5412
Qy 3287 TTCTCCCTTCGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTGGG 3346
Db 5411 TTCTCCCTTCGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTGGG 5352
Qy 3347 TGTAGTCTGCTCCAGCTGGGCTGTGTGACGAAACCCCGCTTCAGCCGACCGCT 3406
Db 5351 TGTAGTCTGCTCCAGCTGGGCTGTGTGACGAAACCCCGCTTCAGCCGACCGCT 5292
Qy 3407 GCGCTTATCCGCTAACTATCTGTTGAGTCCAAACCGGTAAAGACACACTTATCGCCAC 3466
Db 5291 GCGCTTATCCGCTAACTATCTGTTGAGTCCAAACCGGTAAAGACACACTTATCGCCAC 5232
Qy 3467 TGGCAGACCCACTCGTAAACAGATTAGCAGCGAGGTATGTAGGCGGTGCTACAGAT 3526
Db 5231 TGGCAGACCCACTCGTAAACAGATTAGCAGCGAGGTATGTAGGCGGTGCTACAGAT 5172
Qy 3527 TCTTGAAGTGGTGGCTTAACCTACGCTACACTAGAGAACAGTATTTGGTATCTCGGCTC 3586
Db 5171 TCTTGAAGTGGTGGCTTAACCTACGCTACACTAGAGAACAGTATTTGGTATCTCGGCTC 5112
Qy 3587 TGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTTGATCCGGCAAAACCA 3646
Db 5111 TGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTTGATCCGGCAAAACCA 5052
Qy 3647 CGCTGGTAGCGGTGTTTTTTTGTTCACAGCAGCAGATTACCGCGCAGAAAAAGGAT 3706
Db 5051 CGCTGGTAGCGGTGTTTTTTTGTTCACAGCAGCAGATTACCGCGCGCCGAGTGAGG 4992
Qy 3707 CTCACAGATCTTTTGTATCTTTTCTACGGGTCTCAGGCTC-----AGTGGAAAG 3757
Db 4991 GGTGTGGGCTCTTTTATTTAGCTCGGGAGCAGAAAGCGCGAACAGAGAGAGCG 4932
Qy 3758 AAACTCAAGTTAAGGATTTTGGTCAAGATTATCGTCGACCAAGGCGCATCGTGC 3817
Db 4931 AACTGATTGGTTAGTTCAAATAAGGCACAGGGTCAATTCAGTCTCTTGGGCAACCTGGA 4872
Qy 3818 CTC-----CCCACTCTCGAGTTCGGGGGATGATGCGCGGATAGCGGCTG 3864
Db 4871 AACATCTGATGTTCTCTAGAACTGCTGAGGCTGGACCGCATCTGGGACCATCTGTT 4812
Qy 3865 CTGGTTTCTTGGATGCGGACGATTTC-----ACTGCGGTAGAACCTCCCGGAG 3914
Db 4811 CTTGGGCTTGAAGCGGGGAGAACTGTTACACAGATATCTGTTTGGCCCATATTC 4752
Qy 3915 GTCGTCCAGCTCAGCAGCAGCTGAACCACTCGCGAGGGATCGAGCCCGGGTGGG 3974
Db 4751 GCTGTTCCATCTGTTCTTGGCCCTGAGCGCGGAGAACTGCTTACCAAGATCTCTG 4692
Qy 3975 GAAGAACTCCAGCATGAGATCCCGCGCTGGAGATCATCCAGCGGCTCCCGGAAAC 4034
Db 4691 TTTGGCCCATATTCAGCTGTTCCATCTGTTCTCTGACCTTGAATCTGAATCTCTATCTC 4632
Qy 4035 GATTCCGAAGCCCAACCTTTTCATAGAGCGGGGTGGATCGAAATCTCG----- 4085
Db 4631 AGTTATGTTATTTTCCATGCTTTGCAAAATGGCGGTACTTAAGCTAGTTGCCAAACCTA 4572

QY	4086	-----TGATGGCAGGTTGGGCGTGCCTTGGTTCGGTCAATT	4119
Db	4571	CAGGTGGGTCCTTTCATTCCCCCTTTTCTGGAGGTTGGGCGTGCCTTGGTTCGGTCAATT	4512
QY	4120	TCGAACCCAGAGTCCCGCTCAGAAGACTCGTCAAGAAGCGATAGAAGCGGATGCGCT	4179
Db	4511	TCGAACCCAGAGTCCCGCTCAGAAGACTCGTCAAGAAGCGATAGAAGCGGATGCGCT	4452
QY	4180	GCGAATCGGAGCGGCGATACCGTAAAGCAGAGGAAGCGGTGAGCCCATTTGCCGCCAA	4239
Db	4451	GCGAATCGGAGCGGCGATACCGTAAAGCAGAGGAAGCGGTGAGCCCATTTGCCGCCAA	4392
QY	4240	GCTCTTCAGCAATATCACGGGTAGCCAAAGCTATGTCCTGATAGCGGTCCGCCACACCA	4299
Db	4391	GCTCTTCAGCAATATCACGGGTAGCCAAAGCTATGTCCTGATAGCGGTCCGCCACACCA	4332
QY	4300	GCGGCGCACAGTCGATGAATCCAGAAAAGCGGCCATTTTCCACCATGATATTGGGCAAGC	4359
Db	4331	GCGGCGCACAGTCGATGAATCCAGAAAAGCGGCCATTTTCCACCATGATATTGGGCAAGC	4272
QY	4360	AGGCATCGCATCGGTTCAGACAGAGATCCTCGCGGTCGGGCATGCGGCTTGAGCGCTGG	4419
Db	4271	AGGCATCGCATCGGTTCAGACAGAGATCCTCGCGGTCGGGCATGCGGCTTGAGCGCTGG	4212
QY	4420	CGAACAGTTCCGGTGGCGGAGCGCCCTGATGCTCTTCGTCCAGATCATCCTGATGACAAA	4479
Db	4211	CGAACAGTTCCGGTGGCGGAGCGCCCTGATGCTCTTCGTCCAGATCATCCTGATGACAAA	4152
QY	4480	GACCGGTTCCATCGAGTAGTCTCGTCTCGATGCGATGTTTCGGTTCGGTTCGAAATG	4539
Db	4151	GACCGGTTCCATCGAGTAGTCTCGTCTCGATGCGATGTTTCGGTTCGGTTCGAAATG	4092
QY	4540	GGCAGGTAGCCGATCAAGCGTATGCAGCGCGCGCATTTGCATCAGCCATGATGATACATT	4599
Db	4091	GGCAGGTAGCCGATCAAGCGTATGCAGCGCGCGCATTTGCATCAGCCATGATGATACATT	4032
QY	4600	TCTCGGCAGAGCAAGGTGAGATGACAGAGATCCTGCCCCGGCATTTGCCCAATAGCA	4659
Db	4031	TCTCGGCAGAGCAAGGTGAGATGACAGAGATCCTGCCCCGGCATTTGCCCAATAGCA	3972
QY	4660	GCCAGTCCCTTCGCGCTTCAGTGCACACGTCGAGACACAGTCGCGCAAGGAAGCCCGTCTG	4719
Db	3971	GCCAGTCCCTTCGCGCTTCAGTGCACACGTCGAGACACAGTCGCGCAAGGAAGCCCGTCTG	3912
QY	4720	TGGCAGCCACGATAGCGCGCTGCCTTCGCTCTGTCAGATTCAITCAGGGCACCGGCACAGGT	4779
Db	3911	TGGCAGCCACGATAGCGCGCTGCCTTCGCTCTGTCAGATTCAITCAGGGCACCGGCACAGGT	3852
QY	4780	CGGTCTTGACAAAAGAACCGGGCGCCCTGCGCTGACAGCCGGAACACCGCGGCATCAG	4839
Db	3851	CGGTCTTGACAAAAGAACCGGGCGCCCTGCGCTGACAGCCGGAACACCGCGGCATCAG	3792
QY	4840	AGCAGCCGATTTGTCGTGTGTCGCCAGTCAAGCCGATAGCCGATAGCCCTCTCCACCCAAAGCGGCGG	4899
Db	3791	AGCAGCCGATTTGTCGTGTGTCGCCAGTCAAGCCGATAGCCGATAGCCCTCTCCACCCAAAGCGGCGG	3732
QY	4900	GAGAACCTGGTGCAATTCATCTTTGTTCAATCATGCGAAACGATCTCTCATCTCTGCTCTTT	4959
Db	3731	GAGAACCTGGTGCAATTCATCTTTGTTCAATCATGCGAAACGATCTCTCATCTCTGCTCTTT	3672
QY	4960	GATCAGATCTTGATCTCCCTCGGCCATCAGATCTCTTGGCGGCAAGAAAGCCATCCAGTTTA	5019
Db	3671	GATCAGATCTTGATCTCCCTCGGCCATCAGATCTCTTGGCGGCAAGAAAGCCATCCAGTTTA	3612
QY	5020	CTTTTGCAGGGCTTCCCAACCTTACACAGAGGCGCCCGAGCTGGCAATTCGCGTTTCGCTTG	5079
Db	3611	CTTTTGCAGGGCTTCCCAACCTTACACAGAGGCGCCCGAGCTGGCAATTCGCGTTTCGCTTG	3552
QY	5080	CTGTCCATAAAACCGCCAGTCTAGCTATCGCCATGTAAAGCCACTGCAAGCTACCTGCT	5139
Db	3551	CTGTCCATAAAACCGCCAGTCTAGCTATCGCCATGTAAAGCCACTGCAAGCTACCTGCT	3492
QY	5140	TTCTCTTTTGGCGCTTGGCTTTTCCCTTGTCCAGATAGCCGATAGCTGACATTCATCCGGG	5199

```

Db      3491  TTCTCTTTGGCTTGGCTTTTCCCTTGTCCAGATAGCCAGTAGCTGACATTTCATCCGG 3432
Qy      5200  GTACGACACCGTTTCTCGGACTCGCTTTCTACGTGTTCGCTTCCCTTTACAGACCCCTTGC 5259
Db      3431  GTACGACACCGTTTCTCGGACTCGCTTCTTACGTGTTCGCTTCCCTTTAGCAGCCCTTGC 3372
Qy      5260  GCCTGAGTGCTTGGCGCAGCGTG 5283
Db      3371  GCCTGAGTGCTTGGCGCAGCGTG 3348

RESULT 4
US-09-380-190A-30/c
; Sequence 30: Application US/09380190A
; Patent No. 5410220
; GENERAL INFORMATION:
; APPLICANT: NATURE TECHNOLOGY CORPORATION, ET AL.
; TITLE OF INVENTION: SELF-ASSEMBLING GENES, VECTORS AND USES
; THEREOF
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MUETING, RAASCH & GEBHARDT, P.A.
; STREET: 119 NORTH FOURTH STREET, SUITE 203
; CITY: MINNEAPOLIS
; STATE: MINNESOTA
; COUNTRY: USA
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/380,190A
; FILING DATE: 26-Aug-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/03918
; FILING DATE: 28-FEB-98
; ATTORNEY/AGENT INFORMATION:
; NAME: MUETING, ANN M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 228.00010201
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-305-1217
; TELEFAX: 612-305-1228
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6561 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 30:
US-09-380-190A-30

```

	Query Match	30.6%	Score 1618.2	DB 4	Length 6561
	Best Local Similarity	86.7%	Pre. No. 0		
	Matches 1911	Conservative 0	Mismatches 23	Indels 269	Gaps 1
QY	3081	GCAAAAGGCAGGAACCGTAAAGGCGCGCTTCTCGCTGGCGTTTTTCCATAGGCTCCGCC			3140
DB	2305	GCAGGTGTACTCGAGGCGGCGCATCTCGGCGTGTCTGGCGTTTTTCCATAGGCTCCGCC			2246
QY	3141	CCCTTGACGAGCATCACAAAATCAACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACT			3200
DB	2245	CCCTTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACT			2186
QY	3201	ATAAGATACCGAGCGGTTTCCCTCGAAGCTCCCTCGTGGCTCTCTGTGTTCGACCCCT			3260
DB	2185	ATAAGATACCGAGCGGTTTCCCTCGAAGCTCCCTCGTGGCTCTCTGTGTTCGACCCCT			2126

3261 GCCGTTACCGGATACCTGTCGGCTTCTCCCTTCGGGAAGCGTGGCGTCTTCTCATAG 3320
Db
2125 GCCGTTACCGGATACCTGTCGGCTTCTCCCTTCGGGAAGCGTGGCGTCTTCTCATG 2066
Qy
3321 CTCAGCGTGTAGGTATCTCAGTTCTGGTGTAGGTCTGCTCCAGCTGGCTGTGTGCA 3380
Db
2065 CTCAGCGTGTAGGTATCTCAGTTCTGGTGTAGGTCTGCTCCAGCTGGCTGTGTGCA 2006
Qy
3381 CGAACCCCGGTTACGCGCGGCTGCGCTTATCCGGTAACCTATCGTCTTCTAGTCCAA 3440
Db
2005 CGAACCCCGGTTACGCGCGGCTGCGCTTATCCGGTAACCTATCGTCTTCTAGTCCAA 1946
Qy
3441 CCCGTAAGACACGATCTATCCGACCTGCGACGACCACTGGTAAACAGGATTAGCAGC 3500
Db
1945 CCCGTAAGACACGATCTATCCGACCTGCGACGACCACTGGTAAACAGGATTAGCAGC 1886
Qy
3501 GAGGTATGTAGCGGTGTACAGAGTCTTCAAGTGGTGGCTTAACCTACGCTACACTAG 3560
Db
1885 GAGGTATGTAGCGGTGTACAGAGTCTTCAAGTGGTGGCTTAACCTACGCTACACTAG 1826
Qy
3561 AAGAACAGATATTGTTGTATCTGCGCTCTGCTGAAGCAGTTACCTTCGGAAAAAGAGTTGG 3620
Db
1825 AAGAACAGATATTGTTGTATCTGCGCTCTGCTGAAGCAGTTACCTTCGGAAAAAGAGTTGG 1766
Qy
3621 TAGCTCTTGATCCGCAACCAACCAACCGCTGGTAGCGTGTGTTTTTTTGTTCGAAGCA 3680
Db
1765 TAGCTCTTGATCCGCAACCAACCAACCGCTGGTAGCGTGTGTTTTTTTGTTCGAAGCA 1706
Qy
3681 GCAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCTTTTGATCTTTCTACGGGGTC 3740
Db
1705 GCAGATTACG----- 1696
Qy
3741 TCAGCTCAGTGGACGAAACCTCAGTTAAGGATTTTGTGCATGAGATTATCGTCAC 3800
Db
1695 ----- 1696
Qy
3801 CAAAGCGCCATCGTGCTCCCACTCCCTCAGTTTCGGGGCATGATGCGCGATAGCC 3860
Db
1695 ----- 1696
Qy
3861 GCTGCTGTTTCTGATGCGACGAGTTTGCATGCGGTAGAACTCCGGAGGTGCTG 3920
Db
1695 ----- 1696
Qy
3921 CAGCCTCAGGACAGCTGAACCACTCGCAGGGGATCGAGCCCGGGTGGCGAAGAA 3980
Db
1695 -----AATTCCGGGGTGGCGAAGAA 1675
Qy
3981 CTCAGCATGAGATCCCGGCTGGAGGATCATCAGCCGGGCTCCCGAAAAAGATTCC 4040
Db
1674 CTCAGCATGAGATCCCGGCTGGAGGATCATCAGCCGGGCTCCCGAAAAAGATTCC 1615
Qy
4041 GAAGCCCAACCTTTCATGAAGCGCGGTGGAATCGAAATCTCGTGTGGCAGGTGGG 4100
Db
1614 GAAGCCCAACCTTTCATGAAGCGCGGTGGAATCGAAATCTCGTGTGGCAGGTGGG 1555
Qy
4101 CGTGCTTGGTGGTCAATTCGAAACCCAGAGTCCGCTCAGAAACTCTGTCGAAGAGG 4160
Db
1554 CGTGCTTGGTGGTCAATTCGAAACCCAGAGTCCGCTCAGAAACTCTGTCGAAGAGG 1495
Qy
4161 CGATGAAGCGGATCGCTGCAATCGGAGCGGGATACCTTAAGCAGCAGGAGCGG 4220
Db
1494 CGATGAAGCGGATCGCTGCAATCGGAGCGGGATACCTTAAGCAGCAGGAGCGG 1435
Qy
4221 TCAGCCCATTCGCGCAAGCTTTCAGCAATATCAAGGATATCAAGGATATCAAGGATATCAAGG 4280
Db
1434 TCAGCCCATTCGCGCAAGCTTTCAGCAATATCAAGGATATCAAGGATATCAAGGATATCAAGG 1375
Qy
4281 TAGCGTCCGCAACCCAGCGGCAACAGTGTGATGATTCGAAGAGCGGCAATTTCC 4340
Db
1374 TAGCGTCCGCAACCCAGCGGCAACAGTGTGATGATTCGAAGAGCGGCAATTTCC 1315
Qy
4341 ACCATGATATTCGCAAGCAGGATCGCCATGCGGTCAACGAGATCTTCGCGCGCGG 4400

1314 ACCATGATATTCGCAAGCAGGATCGCCATGGGTACGAGAGATCTTCGCGTTCGGG 1255
Qy
4401 ATGCGCGCTTCAGCTCGGCGAAGCAGTTTCGGTTCGGCGAGCCCCCTGATGCTTTCGTC 4460
Db
1254 ATGCGCGCTTCAGCTCGGCGAAGCAGTTTCGGTTCGGCGAGCCCCCTGATGCTTTCGTC 1195
Qy
4461 AGATCATCTGATTCGACAGACCGGCTTCATTCGAGTACGTTGCTGCTGATGCGGATG 4520
Db
1194 AGATCATCTGATTCGACAGACCGGCTTCATTCGAGTACGTTGCTGCTGATGCGGATG 1135
Qy
4521 TTCGCTTGTGTCGCAATTCGGCAGGTAGCCGATCAAGCTATGAGCGCGCGCATTCGA 4580
Db
1134 TTCGCTTGTGTCGCAATTCGGCAGGTAGCCGATCAAGCTATGAGCGCGCGCATTCGA 1075
Qy
4581 TCAGCCATGATGGATACCTTTCGCGCAGGAGCAAGGTAGATGACAGAGATCTTCGCCCC 4640
Db
1074 TCAGCCATGATGGATACCTTTCGCGCAGGAGCAAGGTAGATGACAGAGATCTTCGCCCC 1015
Qy
4641 GGCACCTTCGCCCAATAGCAGCGAGTCCCTTCCCGCTTCAGTGACAACTCGAGCAAGCT 4700
Db
1014 GGCACCTTCGCCCAATAGCAGCGAGTCCCTTCCCGCTTCAGTGACAACTCGAGCAAGCT 955
Qy
4701 GCGCAAGGAAACCCCGTTCGTCAGGAGCAAGTACGCGGCTGCTTCGTCGAGTTCA 4760
Db
954 GCGCAAGGAAACCCCGTTCGTCAGGAGCAAGTACGCGGCTGCTTCGTCGAGTTCA 895
Qy
4761 TTCAGGCGACCGGACAGGTCGCTTCGACAAAGAAACCGGCGCGCTTCGCTGACAGC 4820
Db
894 TTCAGGCGACCGGACAGGTCGCTTCGACAAAGAAACCGGCGCGCTTCGCTGACAGC 835
Qy
4821 CGGAAACACCGCGGATTCAGAGCAGCGATTTGCTGTTGTCGAGTATAGCGGATAGC 4880
Db
834 CGGAAACACCGCGGATTCAGAGCAGCGATTTGCTGTTGTCGAGTATAGCGGATAGC 775
Qy
4881 CTCTCCACCAAGCGCGGAGAACTCGTGCAATTCATCTGTTCAATCATGCGAAG 4940
Db
774 CTCTCCACCAAGCGCGGAGAACTCGTGCAATTCATCTGTTCAATCATGCGAAG 715
Qy
4941 GATCCTCATCTGCTCTTCCTGATCAGATCTTGTATCCCTTCGCGCATCAGATCTTCGCGC 5000
Db
714 GATCCTCATCTGCTCTTCCTGATCAGATCTTGTATCCCTTCGCGCATCAGATCTTCGCGC 655
Qy
5001 AAGAAAGCATCAGTTTACTTTCAGGCTTCCCAACCTTACCAGAGGCGCGCCAGCT 5060
Db
654 AAGAAAGCATCAGTTTACTTTCAGGCTTCCCAACCTTACCAGAGGCGCGCCAGCT 595
Qy
5061 GGCATTTCCGCTTCGCTTCGTCATATAAACCGCCAGCTTAGCTATCGCCATGTAAGC 5120
Db
594 GGCATTTCCGCTTCGCTTCGTCATATAAACCGCCAGCTTAGCTATCGCCATGTAAGC 535
Qy
5121 CCACTGCAAGCTACCTGCTTCTCTTTGCGCTTGGTTTCCCTGTCAGATAGCCAG 5180
Db
534 CCACTGCAAGCTACCTGCTTCTCTTTGCGCTTGGTTTCCCTGTCAGATAGCCAG 475
Qy
5181 TAGCTGACATTCATCCGGGTGAGCACTTCCTGCGGACTGCTTTCAGCTGTTCCGC 5240
Db
474 TAGCTGACATTCATCCGGGTGAGCACTTCCTGCGGACTGCTTTCAGCTGTTCCGC 415
Qy
5241 TTCCTTTAGCAGCCCTTGGCCCTGAGTGTTCGCGCAGCGTG 5283
Db
414 TTCCTTTAGCAGCCCTTGGCCCTGAGTGTTCGCGCAGCGTG 372

RESULT 5
US-08-472-809B-8
; Sequence 8, Application US/08472809B
; Patent No. 5925564
; GENERAL INFORMATION:
; APPLICANT: Schwartz, Robert J.
; APPLICANT: Demayo, Franco J.
; APPLICANT: O'Malley, Bert W.
; TITLE OF INVENTION: Expression Vector Systems and

TITLE OF INVENTION: Method of Use
 NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: 633 West Fifth Street
 STREET: Suite 4700
 CITY: Los Angeles
 STATE: California
 COUNTRY: U.S.A.
 ZIP: 90071-2066
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 MEDIUM TYPE: Storage
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: IBM P.C. DOS 5.0
 SOFTWARE: Word Perfect 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/472,809B
 FILING DATE: June 7, 1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/209,846
 FILING DATE: March 9, 1994
 APPLICATION NUMBER: 07/789,919
 FILING DATE: No. 5925564member 6, 1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Warburg, Richard J.
 REGISTRATION NUMBER: 32,327
 REFERENCE/DOCKET NUMBER: 214/212
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEX: 67-3510
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5707 bases
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-08-472-809B-8

Query Match 30.3%; Score 1602.2; DB 2; Length 5707;
 Best Local Similarity 82.8%; Pred. No. 0;
 Matches 2021; Conservative 0; Mismatches 28; Indels 392; Gaps 1;

Qy	2706	CTAGACGTAATCATGGTCATAGCTGTTTCTGTGTGAATGTTATCCGCTCACATTC	2765
Db	3514	CTTGGCGTAATCATGGTCATAGCTGTTTCTGTGTGAATGTTATCCGCTCACATTC	3573
Qy	2766	ACACACATACGACCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCCTAATGAGTGAGCTA	2825
Db	3574	ACACACATACGACCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCCTAATGAGTGAGCTA	3633
Qy	2826	ACTCACATTAATTCGGTTGGCTCAGTCGCCGCTTCCAGTCGGGAACCTGTCGTGCCA	2885
Db	3634	ACTCACATTAATTCGGTTGGCTCAGTCGCCGCTTCCAGTCGGGAACCTGTCGTGCCA	3693
Qy	2886	GCTGCATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTGCGTATTGGGCGCTCTTC	2945
Db	3694	GCTGCATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTGCGTATTGGGCGCTCTTC	3753
Qy	2946	CGTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGGTTCGGTCGGCGAGCGGTATCAGC	3005
Db	3754	CGTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGGTTCGGTCGGCGAGCGGTATCAGC	3813
Qy	3006	TCACCTCAAGGCGGTATACGGTTATCCACAGATACAGGGGATTAACGACGAAAGACAT	3065
Db	3814	TCACCTCAAGGCGGTATACGGTTATCCACAGATACAGGGGATTAACGACGAAAGACAT	3873
Qy	3066	GTGAGCAAAAGCCAGCAAAAGGCGAGAACCGTAAAGGCCCGCTTGTGCGGTTTTT	3125
Db	3874	GTGAGCAAAAGCCAGCAAAAGGCGAGAACCGTAAAGGCCCGCTTGTGCGGTTTTT	3933

Qy	3126	CCATAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGTGGCG	3185
Db	3934	CCATAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGTGGCG	3993
Qy	3186	AAACCCGACAGGACTATAAGATACACAGGCGTTCCTCCCTGGAAGCTCCCTCGTGGCTC	3245
Db	3994	AAACCCGACAGGACTATAAGATACACAGGCGTTCCTCCCTGGAAGCTCCCTCGTGGCTC	4053
Qy	3246	TCCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTCGGGAACGCT	3305
Db	4054	TCCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTCGGGAACGCT	4113
Qy	3306	GGCGCTTTCATAGCTCAAGCTAGTATCTCAGTTCGGTGTAGTTCGCTCCAA	3365
Db	4114	GGCGCTTTCATAGCTCAAGCTAGTATCTCAGTTCGGTGTAGTTCGCTCCAA	4173
Qy	3366	GCTGGCTGTGTGACGAAACCCCGCTTACGCGGACCGCTGCGCTTATCCGTAACCTA	3425
Db	4174	GCTGGCTGTGTGACGAAACCCCGCTTACGCGGACCGCTGCGCTTATCCGTAACCTA	4233
Qy	3426	TCGCTTTAGTCCAAACCGGTAGACAGACTTATCCCACTGCGCAGCAGCCTGTAA	3485
Db	4234	TCGCTTTAGTCCAAACCGGTAGACAGACTTATCCCACTGCGCAGCAGCCTGTAA	4293
Qy	3486	CAGGATTAGCAGAGCAGAGTATGTAGCGGTGTACAGAGTTCCTGAAGTGGTGGCTAA	3545
Db	4294	CAGGATTAGCAGAGCAGAGTATGTAGCGGTGTACAGAGTTCCTGAAGTGGTGGCTAA	4353
Qy	3546	CTACGGCTACACTAGAAAGACAGTATTGGTGTATCTGCGCTCTGTGAAGCAGTACCTT	3605
Db	4354	CTACGGCTACACTAGAAAGACAGTATTGGTGTATCTGCGCTCTGTGAAGCAGTACCTT	4413
Qy	3606	CGGAAAGAGTGTGTAGTCTTTCATCCGCAACAAACACCGCTGTCGCGTGGTT	3665
Db	4414	CGGAAAGAGTGTGTAGTCTTTCATCCGCAACAAACACCGCTGTCGCGTGGTT	4473
Qy	3666	TTTTGTTTGCAGCAGCAGATTAACGCGCAGAAAAAAGGATCTCAAGAGATCTTTGAT	3725
Db	4474	TTTTGTTTGCAGCAGCAGATTAACGCGCAGAAAAAAGGATCTCAAGAGATCTTTGAT	4533
Qy	3726	CTTTTCTACGGGGTCTGACGCTCAGTGAAGCAAACTCAGGTTAAGGATTTTGGTCA	3785
Db	4534	CTTTTCTACGGGGTCTGACGCTCAGTGAAGCAAACTCAGGTTAAGGATTTTGGTCA	4551
Qy	3786	GAGATTATCGTCGACCAAGCGGCATCGTGCCTCCCACTCTCGCAGTTCGGGGCATG	3845
Db	4552	-----	4551
Qy	3846	GATCGCGGATAGCGGCTGCTGTTTCTGATGCCGACGGATTTGCACTCCCGGTAGAA	3905
Db	4552	-----	4551
Qy	3906	CTCCGCGAGGTCGTCACGCTCAGGACAGCTGAACCAACTCGCGAGGGGATCGAGCCC	3965
Db	4552	-----	4551
Qy	3966	GGGGTGGCGAAGAACTCCAGCATGAGATCCCGGCTGGAGGATCATCCAGCGCGCTC	4025
Db	4552	-----	4551
Qy	4026	CGGAAACGATTCGGAAGCCCAACCTTTTCATAGAGCGCGGTTGGAATCGAAATCTCG	4085
Db	4552	-----	4551
Qy	4086	TGATGGCAGGTTGGCGCTCGCTTGTGCTCATTTTCGAACCCCAAGTCCCGCTCAGAG	4145
Db	4552	-----	4561
Qy	4146	AACTCGTCAAGAGCGGATAGAGCGGATGCGCTCGGATCGGAGCGCGGATACCGTAA	4205
Db	4562	AACTCGTCAAGAGCGGATAGAGCGGATGCGCTCGGATCGGAGCGCGGATACCGTAA	4621

QY 4206 AGCAGGAGGAGGCTCAGCCATTCGCGCCCAAGCTCTTCAGCAATATCAGGGTAGCC 4265
DB 4622 AGCAGGAGGAGGCTCAGCCATTCGCGCCCAAGCTCTTCAGCAATATCAGGGTAGCC 4681
QY 4266 AACGCTATGTCCTGATAGCGGTCCGCAACACCCAGCCGCGCCACAGTCAGTGAATCCAGAA 4325
DB 4682 AACGCTATGTCCTGATAGCGGTCCGCAACACCCAGCCGCGCCACAGTCAGTGAATCCAGAA 4741
QY 4326 AAGCGGCCATTTCCACCATGATATTCGCAAGCAGGAGGATCGCCATGGTCCAGCAGGA 4385
DB 4742 AAGCGGCCATTTCCACCATGATATTCGCAAGCAGGAGGATCGCCATGGTCCAGCAGGA 4801
QY 4386 TCCTCGCGTGGGAGATGCGCGCTTGAGCTTGGCGAAACAGTTCGCTGGCGAGCC 4445
DB 4802 TCCTCGCGTGGGAGATGCGCGCTTGAGCTTGGCGAAACAGTTCGCTGGCGAGCC 4861
QY 4446 TGATGCTCTTCGTCAGATCATCTGATCGACCAACCGCTTTCATCCGAGTACGTC 4505
DB 4862 TGATGCTCTTCGTCAGATCATCTGATCGACCAACCGCTTTCATCCGAGTACGTC 4921
QY 4506 CGCTCGATGCGATTTTCGCTTGGTGGATGGAATGGCAGGAGTAGCGGATCAAGCGTATGC 4565
DB 4922 CGCTCGATGCGATTTTCGCTTGGTGGTGGATGGAATGGCAGGAGTAGCGGATCAAGCGTATGC 4981
QY 4566 AGCGCGGATGTCATCAGCCATGATGATCTTCTCGGAGGAGCAAGTGAATGATGAC 4625
DB 4982 AGCGCGGATGTCATCAGCCATGATGATCTTCTCGGAGGAGCAAGTGAATGATGAC 5041
QY 4626 AGGAGATCTCGCCCGGCACTTCGCGCAATAGCAGCCAGTCCCTTCGCTTCAGTGACA 4685
DB 5042 AGGAGATCTCGCCCGGCACTTCGCGCAATAGCAGCCAGTCCCTTCGCTTCAGTGACA 5101
QY 4686 ACCTCGAGCAGCTGCGCAAGCAACCGCGTCGCGGACGACGATAGCGCGGTCGC 4745
DB 5102 ACCTCGAGCAGCTGCGCAAGCAACCGCGTCGCGGACGACGATAGCGCGGTCGC 5161
QY 4746 TCGTCTGAGATTCATTCAGGGACCGGAGAGTCCGTCTTGCAAAAGAAACCGGCGC 4805
DB 5162 TCGTCTGAGATTCATTCAGGGACCGGAGAGTCCGTCTTGCAAAAGAAACCGGCGC 5221
QY 4806 CCTCGCGCTGACAGCGGCAACCGGCGCATCAGAGCAGCGATGCTGTGTGCGCCAG 4865
DB 5222 CCTCGCGCTGACAGCGGCAACCGGCGCATCAGAGCAGCGATGCTGTGTGCGCCAG 5281
QY 4866 TCATAGCCGATAGCTCTCCACCAAGCGCGCGAGAACCTGCGTGAATCCATCTGT 4925
DB 5282 TCATAGCCGATAGCTCTCCACCAAGCGCGCGAGAACCTGCGTGAATCCATCTGT 5341
QY 4926 TCAATCATGCGAAACGATCCTCATCTGCTCTTCATCAGATCTTGATCCCGCTCGGCAT 4985
DB 5342 TCAATCATGCGAAACGATCCTCATCTGCTCTTCATCAGATCTTGATCCCGCTCGGCAT 5401
QY 4986 CAGATCTTGGGCGCAAGAACGATTCAGTTTACTTTGAGGGGTTCCCAACCTTACCA 5045
DB 5402 CAGATCTTGGGCGCAAGAACGATTCAGTTTACTTTGAGGGGTTCCCAACCTTACCA 5461
QY 5046 GAGGCGCGCCACGCTGCGCAATTCGCGTTGCTGTGCTTCCATATAACCGCCAGTCTAGC 5105
DB 5462 GAGGCGCGCCACGCTGCGCAATTCGCGTTGCTGTGCTTCCATATAACCGCCAGTCTAGC 5521
QY 5106 TATGCGCATGTAGCCCATGCAAGTACCTGCTTCTCTT 5146
DB 5522 AACTGTTGGAGGGGATCGGTTGGGGCTCTTCGCTATT 5562

RESULT 6

US-08-948-378A-7/c
; Sequence 7, Application US/08948378A
; Patent No. 6013258
; GENERAL INFORMATION:
; APPLICANT: Urban, Robert G.
; APPLICANT: Chicx, Roman M.
; APPLICANT: Collins, Edward J.

APPLICANT: Hedley, Mary Lynn
TITLE OF INVENTION: IMMUNOGENIC PEPTIDES FROM
THE HPV E7 PROTEIN
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: Past-Seq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/948,378A
FILING DATE: 09-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 08191/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 4665 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: circular
MOLECULE TYPE: pBIOTOPHPV
US-08-948-378A-7

Query Match 29.7%; Score 1571.6; DB 3; Length 4665;
Best Local Similarity 86.3%; Pred. No. 0;
Matches 1850; Conservative 0; Mismatches 114; Indels 180; Gaps 3;

QY 3008 ACTCAAGGCGGTATACGGTTATCCAGATCAGGGGATACGCGAGGAACATGT 3067
DB 2624 AATCATGCGCGTATACGGTTATCCAGATCAGGGGATACGCGAGGAACATGT 2565
QY 3068 GAGCAAAAGGCGCAGCAAAAGGCGCAGGAACCGTAAAGGCGGTTGCTGGCGTTTTC 3127
DB 2564 GAGCAAAAGGCGCAGCAAAAGGCGCAGGAACCGTAAAGGCGGTTGCTGGCGTTTTC 2505
QY 3128 ATAGCTCGCGCGCGCTGACGAGCATCAAAAATCGACCTCAAGTCAGAGTGGGAA 3187
DB 2504 ATAGCTCGCGCGCGCTGACGAGCATCAAAAATCGACCTCAAGTCAGAGTGGGAA 2445
QY 3188 ACCCCACAGGACTATAAGATACCAAGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTC 3247
DB 2444 ACCCGACAGGACTATAAGATACCAAGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTC 2385
QY 3248 CTGTTCCGACCGCTCGCGTTACCGGATACCTGTCGCGCTTCTCCCTTCGGAAGCGTG 3307
DB 2384 CTGTTCCGACCGCTCGCGTTACCGGATACCTGTCGCGCTTCTCCCTTCGGAAGCGTG 2325
QY 3308 CGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGCTGTAGGTCTCGTCCGAAGC 3367
DB 2324 CGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGCTGTAGGTCTCGTCCGAAGC 2265
QY 3368 TGGGCTGTGTGCAAGAACCCCGCTTCAGCGGACCGCTCGGCTTATCCGTAATATC 3427
DB 2264 TGGGCTGTGTGCAAGAACCCCGCTTCAGCGGACCGCTCGGCTTATCCGTAATATC 2205
QY 3428 GTCTTGTAGTCCAAACCCCGCTTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACA 3487

NAME: Fraser, Janis K.
 REGISTRATION NUMBER: 34, 819
 REFERENCE/POCKET NUMBER: 08191/004002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-542-5070
 TELEFAX: 617-543-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4665 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA
 US-09-169-425C-7

Query Match 29.7%; Score 1571.6; DB 3; Length 4665;

Best Local Similarity 86.3%; Pred. No. 0;
 Matches 1850; Conservative 0; Mismatches 114; Indels 180; Gaps 3;

QY	3008	ACTCAAGGCGGTAAATACGGTTATCCACAGATCAGGGGATACGCGAGGAAAGAACATGT	3067
DB	2624	AATGATCGCGGTAAATACGGTTATCCACAGATCAGGGGATACGCGAGGAAAGAACATGT	2565
QY	3058	GAGCAAAAGGCGAGCAAAAGGCGGACCGTAAAGGCGCGTGTGCGGTTTTTCC	3127
DB	2564	GAGCAAAAGGCGAGCAAAAGGCGGACCGTAAAGGCGCGTGTGCGGTTTTTCC	2505
QY	3128	ATAGGCTCGCGCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGTGGCGAA	3187
DB	2504	ATAGGCTCGCGCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGTGGCGAA	2445
QY	3188	ACCGACAGGACTATAAGATACGAGGCGTTTCCCTCGGAAGCTCCCTCGTGGCTCTC	3247
DB	2444	ACCGACAGGACTATAAGATACGAGGCGTTTCCCTCGGAAGCTCCCTCGTGGCTCTC	2385
QY	3248	CTGTTCCGACCTGCGGCTTACCGGATACCTGTCCGCGCTTCTCCCTTCGGGAAAGCGTGG	3307
DB	2384	CTGTTCCGACCTGCGGCTTACCGGATACCTGTCCGCGCTTCTCCCTTCGGGAAAGCGTGG	2325
QY	3308	CGCTTCTCATAGCTACCGCTAGGTATCTCAGTTCGGTGTAGTTCGTTCCGTCGAAGC	3367
DB	2324	CGCTTCTCATAGCTACCGCTAGGTATCTCAGTTCGGTGTAGTTCGTTCCGTCGAAGC	2255
QY	3368	TGGGCTGTGTGACGAAACCCCGCTTCAGCCGACCGCTGCGCTTATCCGGTAACATATC	3427
DB	2264	TGGGCTGTGTGACGAAACCCCGCTTCAGCCGACCGCTGCGCTTATCCGGTAACATATC	2205
QY	3428	GTCTTGAGTCCAAACCGGTAAGACACGATTATCGCCACTGGCAGCAGCCACTGGTAACA	3487
DB	2204	GTCTTGAGTCCAAACCGGTAAGACACGATTATCGCCACTGGCAGCAGCCACTGGTAACA	2145
QY	3488	GGATTAGCAGGCGAGGTATGAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCCTAACT	3547
DB	2144	GGATTAGCAGGCGAGGTATGAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCCTAACT	2085
QY	3548	ACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGTGGAAGCAGTTACCTTCG	3607
DB	2084	ACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGTGGAAGCAGTTACCTTCG	2025
QY	3608	GAAAAGAGTTGTAGCTCTTGATCCGCGCAACCAACACCGCTGTAGCGGTGGTTTTT	3667
DB	2024	GAAAAGAGTTGTAGCTCTTGATCCGCGCAACCAACACCGCTGTAGCGGTGGTTTTT	1965
QY	3668	TTGTTTGAAGCAGCAGATTTACCGCGAGAAAAAAGGATCTCAAGAGATCTTTGATCT	3727
DB	1964	TTGTTTGAAGCAGCAGATTTACCGCGAGAAAAAAGGATCTCAAGAGATCTTTGATCT	1905
QY	3728	TTTCTACGGGCTGTAGCTCAGTGGAAACCAACTCAGTTAAGGATTTGGTCAATGA	3787
DB	1904	TTTCTACGGGCTGTAGCTCAGTGGAAACCAACTCAGTTAAGGATTTGGTCAATGA	1845
QY	3788	GATTATCGTCGACCAAGCGGCCATCGTGCT-----	3819

DB	1844	GATTATCAAAAGGATCTTTCACCTAGATCTTTAAATATAAAATGAAGTTTAAATCAA	1785
QY	3820	-----CCCACTCTCTCAGTTTCGGG	3840
DB	1784	TCCTAAGTATATATAGTAACTGAGGCTATGGCAGGCGCTTCCCGCCGACGTTGGCTG	1725
QY	3841	GCATGATGCGCGGATAGCCGCTGTGTTCCTGATGCCAGCGATTTGCATCTGCCG	3900
DB	1724	CGAGCCCTGGGCTTTCACCCGAACTTGGGGGTGGGGTGGGAAAGAAAGAACCGCGG	1665
QY	3901	-----TAGAACTCCGCGAGTCTGTCAGCCTCAGGCGAGCAGTGAACCAAC	3946
DB	1664	CGTATTGGCCCAATGGGCTCTCGTGGGTATCGACAGTGCAGCCCTGGGACGAA	1605
QY	3947	TCGCGAGGGGATCGA-----	3961
DB	1604	CCCCCGGTTTATGAACAAACGACCAACACCGTGCCTTTATTCTGTCTTTTATTGCGG	1545
QY	3962	-----GCCCGG	3967
DB	1544	TCATAGCGCGGTTCTTCCGCTATGTCTCTCCGTGTTTCAGTTAGCTTCCCTTAG	1485
QY	3968	GGTGGCGAAGAACTCCAGCATAGATCCCGCGCTGGAGGATCATCCAGCGCGCTGCC	4027
DB	1484	GGTGGCGAAGAACTCCAGCATAGATCCCGCGCTGGAGGATCATCCAGCGCGCTGCC	1425
QY	4028	GGAAACGATTCGGAAGCCCAACCTTCATAGAAGCGCGCTGGAATCGAAATCTCGTG	4087
DB	1424	GGAAACGATTCGGAAGCCCAACCTTCATAGAAGCGCGCTGGAATCGAAATCTCGTG	1365
QY	4088	ATGGCAGGTTGGCGCTCGCTTCGTCATTTCGAACCCCAAGAGTCCGCTCAGAAGAA	4147
DB	1364	ATGGCAGGTTGGCGCTCGCTTCGTCATTTCGAACCCCAAGAGTCCGCTCAGAAGAA	1305
QY	4148	CTGCTCAGAAGCGGATAGAAGCGATGGCTCGGAATCGGAGCGGCGATACCGTAAAG	4207
DB	1304	CTGCTCAGAAGCGGATAGAAGCGATGGCTCGGAATCGGAGCGGCGATACCGTAAAG	1245
QY	4208	CACGAGAAAGCGGTACGCGCATTCGCGCGCAAGAGCTTTCAGCAATATACAGGATGCCAA	4267
DB	1244	CACGAGAAAGCGGTACGCGCATTCGCGCGCAAGAGCTTTCAGCAATATACAGGATGCCAA	1185
QY	4268	CGCTATGCTGTAGTAGCGGTCCGCGACACCGCGCGGCGACAGTCGATGAATCCAGAAA	4327
DB	1184	CGCTATGCTGTAGTAGCGGTCCGCGACACCGCGCGGCGACAGTCGATGAATCCAGAAA	1125
QY	4328	GCGCCCATTTTCCACCATGATATTCGGCAAGCAGGCGATCCCATGGGTACGACGAGATC	4387
DB	1124	GCGCCCATTTTCCACCATGATATTCGGCAAGCAGGCGATCCCATGGGTACGACGAGATC	1065
QY	4388	CTCGCGGTGGGCGATCGCGCTTGAAGCTTGGCGAAACAGTTGGCTGGCGCGAGCCCTG	4447
DB	1064	CTCGCGGTGGGCGATCGCGCTTGAAGCTTGGCGAAACAGTTGGCTGGCGCGAGCCCTG	1005
QY	4448	ATGCTCTTCGTCAGATCATCTGATCGCAAGACCGGCTTCCATCCGAGTACGTCCTCG	4507
DB	1004	ATGCTCTTCGTCAGATCATCTGATCGCAAGACCGGCTTCCATCCGAGTACGTCCTCG	945
QY	4508	CTCGATCGGATGTTTCGCTTGGTGGTGAATGGGCGAGGTAGCCGATCAAGCGTATGCAG	4567
DB	944	CTCGATCGGATGTTTCGCTTGGTGGTGAATGGGCGAGGTAGCCGATCAAGCGTATGCAG	885
QY	4568	CCGCGCATTCGATCAGCCATGATGATCTTCTCGGAGGAGCAAGGTGAGATGACAG	4627
DB	884	CCGCGCATTCGATCAGCCATGATGATCTTCTCGGAGGAGCAAGGTGAGATGACAG	825
QY	4628	GAGATCTGCCCGGCACTTCGCCCAATAGACCGCATCCCTTCCCGCTTCAGTGAAC	4687
DB	824	GAGATCTGCCCGGCACTTCGCCCAATAGACCGCATCCCTTCCCGCTTCAGTGAAC	765
QY	4688	GTCAGACACAGCTGCGCAAGAACCGCGCTGGCGACGACGATAGCGCGCTGCTC	4747

764	Db	GTGAGCAGCAGCTGCGGACGAGAACGCCCTGCTGTGGCAGCCACGATAGCCGCGTGCCTC	705
4748	Qy	GCTCTGCAGTTTCATTTCAGGGCCACCGCAGCTCGTCTTTGACAAAGAAGAACCGGCGGCC	4807
704	Db	GCTTTGCAGTTTCATTTCAGGGCCACCGCAGCTCGTCTTTGACAAAGAACCGGCGGCC	645
4808	Qy	CTGCGCTGACAGCCGGAAACAGGGGGCATCAGAGCAGCGGATGTCTGTGTGTGCCAGTC	4867
644	Db	CTGCGCTGACAGCCGGAAACAGGGGGCATCAGAGCAGCGGATGTCTGTGTGTGCCAGTC	585
4868	Qy	ATAGCCGAATAGCCTCTCCACCCCAAGCGGCCGGAGAACTCGCTGCATCATCTGTTTC	4927
584	Db	ATAGCCGAATAGCCTCTCCACCCCAAGCGGCCGGAGAACTCGCTGCATCATCTGTTTC	525
4928	Qy	AATCATGCGAAACGATCCTCATCTGTCTCTTGTATCAGATCTTG	4971
524	Db	AATCATGCGAAACGATCCTCATCTGTCTCTTGTATCAGATCTTG	481

RESULT 8

```

US-09-759-960-7/c
? Sequence 7, Application US/09759960
? Patent No. 6582704
? GENERAL INFORMATION:
? APPLICANT: Urban, Robert G.
? APPLICANT: Chiciz, Roman M.
? APPLICANT: Collins, Edward J.
? APPLICANT: Hedley, Mary Lynn
? TITLE OF INVENTION: IMMUNOGENIC PEPTIDES FROM THE HPV E7
? TITLE OF INVENTION: PROTEIN
? NUMBER OF SEQUENCES: 33
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Fish & Richardson, P.C.
? STREET: 225 Franklin Street
? CITY: Boston
? STATE: MA
? COUNTRY: US
? ZIP: 02110-2804
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: Windows95
? SOFTWARE: FASTSEQ for Windows Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09759,960

```

Query Match	29.7%	Score	1571.6	DB	4	Length	4665
Best Local Similarity	86.3%	Prod. NO.	0				
Matches 1850	Conservative	0	Mismatches	114	Indels	180	Gaps
3008	ACTCAAGGGCGGTAATACGGTTATCCACAGAAATCAGGGGATAACCCAGGAAAGAAATGT						3067
2624	AATGCATGGGGGTAAATACGGTTATCCACAGAAATCAGGGGATAACCGCAGGAAAGAAATGT						2565

[illegible]

QY 3968 GGTGGCGGAGAACTCCAGCATGAGATCCCGCGCTGGAGATCATCCAGCCGCGCTCC 4027
DB 1484 GGTGGCGGAGAACTCCAGCATGAGATCCCGCGCTGGAGATCATCCAGCCGCGCTCC 1425
QY 4028 GGAAGAACGATTCGGAAGCCCACTTTTCATAGAGGGCGCGGTGGAATCGAAATCTCGTG 4087
DB 1424 GGAAGAACGATTCGGAAGCCCACTTTTCATAGAGGGCGCGGTGGAATCGAAATCTCGTG 1365
QY 4088 ATGGCAGGTGGCGCTGGCTGGCTGATTTTGAACCCAGAGTCCCGCTCAGAAGAA 4147
DB 1364 ATGGCAGGTGGCGCTGGCTGGCTGATTTTGAACCCAGAGTCCCGCTCAGAAGAA 1305
QY 4148 CTCGTCAAGAGCGCATAGAGCGGATCGCTGCGAATCGGGAGCGCGCATACCGTAAAG 4207
DB 1304 CTCGTCAAGAGCGCATAGAGCGGATCGCTGCGAATCGGGAGCGCGCATACCGTAAAG 1245
QY 4208 CAGCAGAGCGGTGAGCCCATTCGCGCGCAAGCTCTTCAGCATATACCGGTAGCCAA 4267
DB 1244 CAGCAGAGCGGTGAGCCCATTCGCGCGCAAGCTCTTCAGCATATACCGGTAGCCAA 1185
QY 4268 CGTATGCTGATAGCGGTGCGCCACACCCAGCGCGCCACAGTCGATGATTCAGAA 4327
DB 1184 CGTATGCTGATAGCGGTGCGCCACACCCAGCGCGCCACAGTCGATGATTCAGAA 1125
QY 4328 GCGGCCATTTTCCACCATGATATTCGCAAGCAGGATCGCCATCGGTACAGCAGATC 4387
DB 1124 GCGGCCATTTTCCACCATGATATTCGCAAGCAGGATCGCCATCGGTACAGCAGATC 1065
QY 4388 CTCGCGTGGGATGCGCGCTTGAGCTGCGCAAGCTTCGCTGCGCGAGCGCCCTG 4447
DB 1064 CTCGCGTGGGATGCGCGCTTGAGCTGCGCAAGCTTCGCTGCGCGAGCGCCCTG 1005
QY 4448 ATGCTCTTCTGTCAGATCATCTGATCGACAGAGCGGCTTCCATCCGAGTACGTGCTG 4507
DB 1004 ATGCTCTTCTGTCAGATCATCTGATCGACAGAGCGGCTTCCATCCGAGTACGTGCTG 945
QY 4508 CTCGATGCGATGTTTCTGCTGCTGCTGCAATGGCAGAGTACCGGATCAGCTATGCG 4567
DB 944 CTCGATGCGATGTTTCTGCTGCTGCTGCAATGGCAGAGTACCGGATCAGCTATGCG 885
QY 4568 CCGCGCATTCGATCAGCATGATGATGATGATGATGATGATGATGATGATGATGATG 4627
DB 884 CCGCGCATTCGATCAGCATGATGATGATGATGATGATGATGATGATGATGATGATG 825
QY 4628 GAGATCCTGCGCGGACCTTCGCAATAGCAGCGCTTCCGCTTCCGCTTCCGCTTCC 4687
DB 824 GAGATCCTGCGCGGACCTTCGCAATAGCAGCGCTTCCGCTTCCGCTTCCGCTTCC 765
QY 4688 GTCGAGCAGCTGCGCAAGGAAACGCGCTGCTGCGCAGCAACGATAGCGCGCTGCTC 4747
DB 764 GTCGAGCAGCTGCGCAAGGAAACGCGCTGCTGCGCAGCAACGATAGCGCGCTGCTC 705
QY 4748 GTCCTGCGATTCATTCAGGCGCACCGGACAGGTCGCTTTCGCAAAAGAACCGCGGCG 4807
DB 704 GTCCTGCGATTCATTCAGGCGCACCGGACAGGTCGCTTTCGCAAAAGAACCGCGGCG 645
QY 4808 CTCGCTGACAGCGGAAACAGCGGCGCATCAGAGCGCGGATGCTGCTGCTGCTGCTGCT 4867
DB 644 CTCGCTGACAGCGGAAACAGCGGCGCATCAGAGCGCGGATGCTGCTGCTGCTGCTGCT 585
QY 4868 ATAGCGCAATGAGCTCTCCACCCCAAGCGCGCGGAGAACCTCGCTGCAATCCATCTGTT 4927
DB 584 ATAGCGCAATGAGCTCTCCACCCCAAGCGCGCGGAGAACCTCGCTGCAATCCATCTGTT 525
QY 4928 AATCATGCGAAGCATCT 4971
DB 524 AATCATGCGAAGCATCT 481

RESULT 9

US-09-380-190A-26

; Sequence 26, Application US/09380190A

; Patent No. 6410220

GENERAL INFORMATION:
APPLICANT: NATURE TECHNOLOGY CORPORATION, ET AL.
TITLE OF INVENTION: SELF-ASSEMBLING GENES, VECTORS AND USES THEREOF
NUMBER OF SEQUENCES: 74
CORRESPONDENCE ADDRESS:
ADDRESSES: MUELLING, RAASCH & GEBHARDT, P.A.
STREET: 119 NORTH FOURTH STREET, SUITE 203
CITY: MINNEAPOLIS
STATE: MINNESOTA
COUNTRY: USA
ZIP: 55401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/09/380,190A
FILING DATE: 26-Aug-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US98/03918
FILING DATE: 28-FEB-98
ATTORNEY/AGENT INFORMATION:
NAME: MUELLING, ANN M.
REGISTRATION NUMBER: 33,977
REFERENCE/DOCKET NUMBER: 228,00010201
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-305-1217
TELEFAX: 612-305-1228
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 4518 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-380-190A-26

Query Match 29.7%; Score 1570; DB 4; Length 4518;
Best Local Similarity 86.2%; Pred. No. 0;
Matches 1849; Conservative 0; Mismatches 115; Indels 180; Gaps 3;
QY 3008 ACTCAAGCGGTAAATACGGTTATCCAGATCAGGGGATAACGACGAGGAAGAACATGT 3067
DB 1895 ATGCTATGGCGGTAATACGGTTATCCAGATCAGGGGATAACGACGAGGAAGAACATGT 1954
QY 3068 GAGCAAAAGGCCAGCAAAAGGCCAGCAAAAGGCCAGGTTGCTGGCGTTTTC 3127
DB 1955 GAGCAAAAGGCCAGCAAAAGGCCAGCAAAAGGCCAGGTTGCTGGCGTTTTC 2014
QY 3128 ATAGGCTCCGCCCTCCGAGATCAAAAATCGAGCTCAAGTCAGAGGTGGCGAA 3187
DB 2015 ATAGGCTCCGCCCTCCGAGATCAAAAATCGAGCTCAAGTCAGAGGTGGCGAA 2074
QY 3188 ACCGACAGGACTATAAAGTACAGGCGTTTCCCTCGAAGCTCCCTCGTGGCTCTC 3247
DB 2075 ACCGACAGGACTATAAAGTACAGGCGTTTCCCTCGAAGCTCCCTCGTGGCTCTC 2134
QY 3248 CTGTTCCGACCTCCGCTTACCGGATACCTGCTCGGCTTTTCTCCCTTCGGGAAGCGTGG 3307
DB 2135 CTGTTCCGACCTCCGCTTACCGGATACCTGCTCGGCTTTTCTCCCTTCGGGAAGCGTGG 2194
QY 3308 CGCTTCTCATAGCTACCGCTAGGTATCTCAGTTCGGTTCAGTTCGCTCCAGC 3367
DB 2195 CGCTTCTCATAGCTACCGCTAGGTATCTCAGTTCGGTTCAGTTCGCTCCAGC 2254
QY 3368 TGGGCTGTGTGACGAACCCCGCTTACGCCCGACCGCTGCGCTTATCCGTAATATC 3427
DB 2255 TGGGCTGTGTGACGAACCCCGCTTACGCCCGACCGCTGCGCTTATCCGTAATATC 2314

QY 3428 GTCTTGGTCCACCGGTAAGACACGACTTATCCGCACTGCGCAGCCACTGGTAACA 3487
 Db 2315 GTCTTGGTCCACCGGTAAGACACGACTTATCCGCACTGCGCAGCCACTGGTAACA 2374
 QY 3488 GGATTAGCAGAGCGAGTATAGAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAACT 3547
 Db 2375 GGATTAGCAGAGCGAGTATAGAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAACT 2434
 QY 3548 ACGGCTACACTAGAGAAAGATTTGGTATCTCGGCTCTGCTGAAGCCAGTACCTTCG 3607
 Db 2435 ACGGCTACACTAGAGAAAGATTTGGTATCTCGGCTCTGCTGAAGCCAGTACCTTCG 2494
 QY 3608 GAAAAAGAGTGTGACTCTTGAATCGGCAAAACAAACACCGCTGTGTAGCGGTGGTTTT 3667
 Db 2495 GAAAAAGAGTGTGACTCTTGAATCGGCAAAACAAACACCGCTGTGTAGCGGTGGTTTT 2554
 QY 3668 TTGTTTGCAGACAGAGATTTAGCGCAGAAAAAGAGATCTCAAGAGATCTTTGATCT 3727
 Db 2555 TTGTTTGCAGACAGAGATTTAGCGCAGAAAAAGAGATCTCAAGAGATCTTTGATCT 2614
 QY 3728 TTTCTACGGGCTGTGAGCTCAGTGAAGCAAACTCAGTTAAGGATTTTGTCTATGA 3787
 Db 2615 TTTCTACGGGCTGTGAGCTCAGTGAAGCAAACTCAGTTAAGGATTTTGTCTATGA 2674
 QY 3788 GATTATCGTGCAGCAAGCGGCATCGTGCCT- 3819
 Db 2675 GATTATCAAAAGGATCTTCACTAGATCTTTTAAATTAAGTTTTAAATCAA 2734
 QY 3820 - 3840
 Db 2735 TCTAAAGTATATAGTAACTGAGGCTATGGCAGGCGCTTCCGCGCGAGCTTGGCTG 2794
 QY 3841 GCATGATGCGCGATAGCGCTGTCTGTTCTCGATCGCGACGATTTGCATCTGCGG 3900
 Db 2795 CGAGCCCTGGCCTTCAACCGAATTTGGGGGTGGGGTGGGAAAGAAAGAAAGCGGG 2854
 QY 3901 - 3946
 Db 2855 CGTATTGGCCCCAATGGGTCTCGTGGGGTATGACACAGTGTCCAGCCCTGGGACCGAA 2914
 QY 3947 TCAGAGAGGATCGA- 3961
 Db 2915 CCGCGCTTTATGAACAAAGACCAACACCGTGGTGTATTTCTGCTTTTATTTGCGG 2974
 QY 3962 - 3967
 Db 2975 TCAATAGCGGGTCTCTCGGTATTTGCTCTCGGTGTTTCAATAGCTTCCCTCCCTAG 3034
 QY 3968 GGTGGCGAAGAACTCCAGCATGAGATCCCGCGCTGGAGGATCATCCAGCGCGCTCC 4027
 Db 3035 GGTGGCGAAGAACTCCAGCATGAGATCCCGCGCTGGAGGATCATCCAGCGCGCTCC 3094
 QY 4028 GGAAGATTCGAAGCCCACTTTCTAGAGCGCGGTGGAATCGAAATCTGCTG 4087
 Db 3095 GGAAGATTCGAAGCCCACTTTCTAGAGCGCGGTGGAATCGAAATCTGCTG 3154
 QY 4088 ATGSCAGTTGGGCTGCTGTGCTGCTCATTTGAAACCCCGAGAGTCCCGCTCAGAGAA 4147
 Db 3155 ATGSCAGTTGGGCTGCTGTGCTGCTCATTTGAAACCCCGAGAGTCCCGCTCAGAGAA 3214
 QY 4148 CTCTCAGAGAGCGATAGAGGATGCTGCTGCAATCGGAGCGGCGATACCGTAAAG 4207
 Db 3215 CTCTCAGAGAGCGATAGAGGATGCTGCTGCAATCGGAGCGGCGATACCGTAAAG 3274
 QY 4208 CACAGAGAGCGGTACGCCATTTCCGCGCAAGCTTTCCAGCAATATATACCGGTAGCCAA 4267
 Db 3275 CACAGAGAGCGGTACGCCATTTCCGCGCAAGCTTTCCAGCAATATATACCGGTAGCCAA 3334
 QY 4268 CGCTATGCTCTGATAGAGGTCGCGCAACCCAGCGCGGCACAGTGTGATGATCCAGAAA 4327
 Db 3335 CGCTATGCTCTGATAGAGGTCGCGCAACCCAGCGCGGCACAGTGTGATGATCCAGAAA 3394
 QY 4328 GCGGCCATTTTCCACCATATGATATTCGGCAAGAGCGCATCGCCATGGGTCAAGCAGATC 4387

Db 3395 GCGGCCATTTTCCACCATGATATTTCCGCAAGCAGGATCGCCATGGGTCAAGCAGATC 3454
 QY 4388 CTCGCGCTGGGATGCGCGCTTGAGCTTGGCGAGACAGTTCGGCTGGCGAGCCCTG 4447
 Db 3455 CTCGCGCTGGGATGCGCGCTTGAGCTTGGCGAGACAGTTCGGCTGGCGAGCCCTG 3514
 QY 4448 ATGCTCTTCTGCTCAGATCATCTCTGATCGAACAAGCCGCTTCCATCCGAGTACGTGCTG 4507
 Db 3515 ATGCTCTTCTGCTCAGATCATCTCTGATCGAACAAGCCGCTTCCATCCGAGTACGTGCTG 3574
 QY 4508 CTCGATGAGTGTTCGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4567
 Db 3575 CTCGATGAGTGTTCGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3634
 QY 4568 CCGCGCATTCATCAGCAGCATGATGGATCTTTCTCGGCAAGGAGCAAGGTGAGATGACAG 4627
 Db 3635 CCGCGCATTCATCAGCAGCATGATGGATCTTTCTCGGCAAGGAGCAAGGTGAGATGACAG 3694
 QY 4628 GAGATCTTCCCGCGCATCTCGCCCAATAGCAGCAGTCCCTTCCGCTTCACTGACAAAC 4687
 Db 3695 GAGATCTTCCCGCGCATCTCGCCCAATAGCAGCAGTCCCTTCCGCTTCACTGACAAAC 3754
 QY 4688 GTGAGCAAGCTGCGCAAGAAACCGCGCTGCTGGCCAGCAGATAGCGCGCTGCTC 4747
 Db 3755 GTGAGCAAGCTGCGCAAGAAACCGCGCTGCTGGCCAGCAGATAGCGCGCTGCTC 3814
 QY 4748 GTCTTGCAGTTTCAATTCAGGCGCACCGGACAGTCTGACAAAAAGAACCCGCGGCC 4807
 Db 3815 GTCTTGCAGTTTCAATTCAGGCGCACCGGACAGTCTGACAAAAAGAACCCGCGGCC 3874
 QY 4808 CTGCGCTCAGCAGCGGAAACCGCGCGCATCAGACAGCCGATTTCTGTTGTCGCGATC 4867
 Db 3875 CTGCGCTCAGCAGCGGAAACCGCGCGCATCAGACAGCCGATTTCTGTTGTCGCGATC 3934
 QY 4868 ATAGCGGATAGCTTCTCCACCCAGCGCGCGGAGAACCTGCGTGCATCATCTTCTTC 4927
 Db 3935 ATAGCGGATAGCTTCTCCACCCAGCGCGCGGAGAACCTGCGTGCATCATCTTCTTC 3994
 QY 4928 AATCATGGAACAGATCTCATCTCTCTCTTGTATCATGATCTTG 4971
 Db 3995 AATCATGGAACAGATCTCATCTCTCTCTTGTATCATGATCTTG 4038

RESULT 10

US-09-533-220A-4/c
 ; Sequence 4, Application US/09533220A
 ; Patent No. 6406908
 ; GENERAL INFORMATION:
 ; APPLICANT: McIntyre, Peter
 ; APPLICANT: James, Iain Fraser
 ; TITLE OF INVENTION: Human Vanilloid Receptor
 ; FILE REFERENCE: 4-30875A
 ; CURRENT APPLICATION NUMBER: US/09/533,220A
 ; PRIOR FILING DATE: 2000-03-23
 ; PRIOR FILING DATE: 1999-03-26
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn Ver. 1.30
 ; SEQ ID NO 4
 ; LENGTH: 4886
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-533-220A-4

Query Match 29.7%; Score 1570; DB 4; Length 4886;
 Best Local Similarity 86.2%; Pred. No. 0;
 Matches 1849; Conservative 0; Mismatches 115; Indels 180; Gaps 3;
 QY 3008 ACTCAAGCGGTAAATACGGTTATCCACAGATCAGGGGATAACGAGGAAGAAACATGT 3067
 Db 4295 AATGATGCGGTAAATACGGTTATCCACAGATCAGGGGATAACGAGGAAGAAACATGT 4236

APPLICANT: ANDERSON, ROBERT J.
 APPLICANT: GRANT, HUGH
 APPLICANT: MACDONALD, IAN D.
 TITLE OF INVENTION: INTERLUKIN-12 FUSION PROTEIN
 NUMBER OF SEQUENCES: 80
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: NIXON & VANDERHYZE P.C.
 STREET: 1100 NORTH GLEBE ROAD
 CITY: ARLINGTON
 STATE: VA
 COUNTRY: USA
 ZIP: 22201
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/751,767A
 FILING DATE: 08-NOV-1996
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: SADOFF, B.J.
 REGISTRATION NUMBER: 36,663
 REFERENCE/DOCKET NUMBER: 117-221
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 7038164091
 TELEFAX: 7038164100
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 6139 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: both
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 3222..4841
 US-08-751-767A-7

Query Match	29.7%; Score 1570; DB 2; Length 6139;
Best Local Similarity	86.2%; Pred. No. 0;
Matches 1843; Conservative	0; Mismatches 115; Indels 180; Gaps 3;
QY	3008 ACTCAAAGGCGGTAAATACGTTATCCACAGAATCAGGGGATAACGACGAGGAAGAAACATGT 3067
Db	2624 AATGCATGGCGGTAAATACGTTATCCACAGATCAGGGATACGCGAGGAGAAACATGT 2565
QY	3068 GAGCAAAAGGCCACAGCAAAAAGGCCAGGAACCGTAAAAAGGCCGGTTCTGGCGTTTTTTC 3127
Db	2564 GAGCAAAAGGCCACAGCAAAAAGGCCAGGAACCGTAAAAAGGCCGGTTCTGGCGTTTTTTC 2505
QY	3128 ATAGGCTTCGCGCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCACAGGTGGCGAA 3187
Db	2504 ATAGGCTTCGCGCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCACAGGTGGCGAA 2445
QY	3188 ACCCGACAGGACTATAAAGATACACAGCGTTTCCCTCGAAGCTCCCTCGTGGCTCTC 3247
Db	2444 ACCCGACAGGACTATAAAGATACACAGCGTTTCCCTCGAAGCTCCCTCGTGGCTCTC 2385
QY	3248 CTGTTTCGACCCCTGCGCCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTGG 3307
Db	2384 CTGTTTCGACCCCTGCGCCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTGG 2325
QY	3308 CGCTTTCTCATAGCTACGCTGTAGGTATCTCAGTTCGGTGTAGGTGGTTTCGCTCCAAGC 3367
Db	2324 CGCTTTCTCATAGCTACGCTGTAGGTATCTCAGTTCGGTGTAGGTGGTTTCGCTCCAAGC 2265
QY	3368 TGGGCTGTGTGACGACACCCCGCTTACGCCGACCGCTGGCCTTATCGGTAACTATC 3427
Db	2264 TGGGCTGTGTGACGACACCCCGCTTACGCCGACCGCTGGCCTTATCGGTAACTATC 2205
QY	3428 GTCTTTGAGTCCAAACCCGGTAAGACACGACTTATCGCACTGGCAGACGCCACTGGTAACA 3487

Db	2204	GTCTTGAGTCCAACCCGGTAAAGACAGACTTATCGCCACTGGCAGCAGCCACTGGTAACA	2145
Qy	3488	GGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTTGAAAGTGTGCGCCTAACT	3547
Db	2144	GGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTTGAAAGTGTGCGCCTAACT	2085
Qy	3548	ACGGCTACACTAGAAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTTACCTTCG	3607
Db	2084	ACGGCTACACTAGAAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTTACCTTCG	2025
Qy	3608	GAAAAAGAGTTGGTAGCTCTTTGATCCGGGCAAAACAAACCCAGCTGGTAGCGGTGGTTTTT	3667
Db	2024	GAAAAAGAGTTGGTAGCTCTTTGATCCGGGCAAAACAAACCCAGCTGGTAGCGGTGGTTTTT	1965
Qy	3668	TTGTTTCCAAAGCAGCAGATTACGGCGAGAAAAAAGATCTCAAGAGATCCCTTTGATCT	3727
Db	1964	TTGTTTCCAAAGCAGCAGATTACGGCGAGAAAAAAGATCTCAAGAGATCCCTTTGATCT	1905
Qy	3728	TTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACCTCACGTTAAGGATTTTGGTCATGA	3787
Db	1904	TTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACCTCACGTTAAGGATTTTGGTCATGA	1845
Qy	3788	GATTATCGTCGACCAAGCGCGCATCTGGCT	3819
Db	1844	GATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTTAAAAATGAAGTTTAAATCAA	1785
Qy	3820	-----CCCCACTCTCGACTTCGGGGCTGCGGCCCGACGTTGGCTG	3840
Db	1784	TCATAAGTATATATGAGTAACCTGAGGCTATGGCAGGGCTGCGGCCCGACGTTGGCTG	1725
Qy	3841	GCATGATGCGCGGATAGCCGCTGCTGGTTTCTGGATGCCGACGATTTGCACTGCGCG	3900
Db	1724	CGAGCCCTGGCCCTTCACCCGAACTTGGGGGTGGGGTGGGGAAGAAAGAACCGGG	1665
Qy	3901	-----TAGAACTCGCGAGGTGCTCGTGGGGTATCGACAGAGTCCGACCCCTGGACCGAA	3946
Db	1664	CGTATGCGCCCAATGGGGTCTCGTGGGGTATCGACAGAGTCCGACCCCTGGACCGAA	1605
Qy	3947	TCGCGAGGGGATCGA-----	3961
Db	1604	CCCCGGTTTATGAACAAACGACCAACCGTGGTTTTATTCTGTCTTTTATTGCGG	1545
Qy	3962	-----GCCCG	3967
Db	1544	TCATAGCGCGGTTCTTCCGGTATGTCTCTTTCGGTGTTCAGTTAGGCTCCCCCTAG	1485
Qy	3968	GGTGGCGAAGAACTCCAGCATAGATCCCGCGCTGGAGGATCATCCAGCCCGCGTCCC	4027
Db	1484	GGTGGCGAAGAACTCCAGCATAGATCCCGCGCTGGAGGATCATCCAGCCCGCGTCCC	1425
Qy	4028	GGAAAAAGATTCCGAAGCCAACTTTTCATAGAAAGCGCGGTGGATCGAAATCTCTGTG	4087
Db	1424	GGAAAAAGATTCCGAAGCCAACTTTTCATAGAAAGCGCGGTGGATCGAAATCTCTGTG	1365
Qy	4088	ATGCGAGTTGGGGTCTGCTGGTGGTCAATTCGAAACCCAGAGTCCCGCTCAGAAGAA	4147
Db	1364	ATGGCAGGTTGGGGTCTGCTGGTGGTCAATTCGAAACCCAGAGTCCCGCTCAGAAGAA	1305
Qy	4148	CTCGTCAAGAGGCGATAGAGGCGATGCGCTGCGAATCGGAGCGCGGATACCGTAAAG	4207
Db	1304	CTCGTCAAGAGGCGATAGAGGCGATGCGCTGCGAATCGGAGCGCGGATACCGTAAAG	1245
Qy	4208	CACGAGGAAGGGTCAGCCCAATTCGCGCCAGCTCTTCAGCATATCACGGGTAGCCAA	4267
Db	1244	CACGAGGAAGGGTCAGCCCAATTCGCGCCAGCTCTTCAGCATATCACGGGTAGCCAA	1185
Qy	4268	CGCTATGTCCTGATAGCGGTCGCGCACACCCAGCCGCGCACAGTCGATGATCCAGAAA	4327
Db	1184	CGCTATGTCCTGATAGCGGTCGCGCACACCCAGCCGCGCACAGTCGATGATCCAGAAA	1125
Qy	4328	GGCGGCAATTTCCACCATGATATTTCGGGACAGCAGGCAATCGCCATGGGTACGACGAGTTC	4387

QY 3793 TCGTCGACCAAGCGCCATCGTGCT 3819
Db 8013 TCAAAAGGATCTTACCTAGATCTTTAAATTAATGAAGTTTAAATCAATCTAA 7954
QY 3820 -----CCCCACTCTCGAGTTTCGGGGGATG 3845
Db 7953 AGTATATAGTAAGTCTAGGCTATGGCAGGCTTGGCCCGCCGACGTTGGCTCGAGC 7894
QY 3846 GATGCGCGGATAGCCCTGCTGCTTCTCTGATGCGACGATTTGCACTGCGG 3900
Db 7893 CTTGGGCTTCAACCGAACTTGGGGGTGGGTGGGAAAGGAAGAACGCGGCGTAT 7834
QY 3901 -----TAGAACTCCGAGGTCGTCCAGCTCAGGACAGAGTGAACCAATCCTCG 3951
Db 7833 TGCGCCCAATGGGCTCTCGTGGGTATCGACAGAGTGCACGCTTGGGACCGAACCGC 7774
QY 3952 AGGGATCGA 3961
Db 7773 CGTTATGAACAAAGACCAACACCGTGGGTTTATTCGTCTTTTATTCGCGTATA 7714
QY 3962 -----GCCCGGGTGG 3972
Db 7713 GCGCGGCTCTTCGCGTATGTCCTTCCGTTTTCAGTTAGCTCCCTAGGGTGG 7654
QY 3973 GCGAAGACTCCAGCATGATCCCGCGCTGGAGATCATCCAGCGCGCTCCGAAA 4032
Db 7653 GCGAAGACTCCAGCATGATCCCGCGCTGGAGATCATCCAGCGCGCTCCGAAA 7594
QY 4033 ACCATTCCGAAGCCCAACCTTTCATAGAAGGGCGGTGGAATCGATCGATGCG 4092
Db 7593 ACCATTCCGAAGCCCAACCTTTCATAGAAGGGCGGTGGAATCGATCGATGCG 7534
QY 4093 AGTTGGGCTCGCTTGGTGGTTCATTTGCAACCCAGAGTCCCGTCAAGAACTCGT 4152
Db 7533 AGTTGGGCTCGCTTGGTGGTTCATTTGCAACCCAGAGTCCCGTCAAGAACTCGT 7474
QY 4153 CRAAGAGGCGATAGAAGCGATCGCTGCGAATCGGGAGCGGATACCGTAAAGCAGA 4212
Db 7473 CAAGAAGCGATAGAAGCGATCGCTGCGAATCGGGAGCGGATACCGTAAAGCAGA 7414
QY 4213 GGAAGCGGTGAGCCCATTCGCCCAAGTCTTTCAGCAATATCACGGTAGCAACGCTA 4272
Db 7413 GGAAGCGGTGAGCCCATTCGCCCAAGTCTTTCAGCAATATCACGGTAGCAACGCTA 7354
QY 4273 TGTCTGTAGTGGTGGCGCACACCCAGCGGCGCAGTGCATGATCCAGAAAGCGC 4332
Db 7353 TGTCTGTAGTGGTGGCGCACACCCAGCGGCGCAGTGCATGATCCAGAAAGCGC 7294
QY 4333 CATTTTCCACCATGATATTCGCAAGCAGGATCGCCATGGGTACGACGAGATCCTCG 4392
Db 7293 CATTTTCCACCATGATATTCGCAAGCAGGATCGCCATGGGTACGACGAGATCCTCG 7234
QY 4393 GTCGGGCTGCGGCTTGGCTGCGGCAAGTTCGGCTGGCGGAGCCCTGATGCT 4452
Db 7233 GTCGGGCTGCGGCTTGGCTGCGGCAAGTTCGGCTGGCGGAGCCCTGATGCT 7174
QY 4453 CTTCTGTCAGATCATCTGATCGACAGACCGGCTTCCATCCAGTACGTCGCTCGA 4512
Db 7173 CTTCTGTCAGATCATCTGATCGACAGACCGGCTTCCATCCAGTACGTCGCTCGA 7114
QY 4513 TCGATTTTTCGTTGGTGGTTCGAATGGGAGGTAGCCGATCAAGCGTAGCAGCGCC 4572
Db 7113 TCGATTTTTCGTTGGTGGTTCGAATGGGAGGTAGCCGATCAAGCGTAGCAGCGCC 7054
QY 4573 GCATTGATCAGCATGATGATGATCTTTCGCGAGGACGCAAGGTGAGATGACAGAGAT 4632
Db 7053 GCATTGATCAGCATGATGATGATCTTTCGCGAGGACGCAAGGTGAGATGACAGAGAT 6994
QY 4633 CTTGCCCGGCACTTCGCCCAATAGCAGCGAGTCCCTTCCCGTTTCAGTGACAACTCGA 4692
Db 6993 CTTGCCCGGCACTTCGCCCAATAGCAGCGAGTCCCTTCCCGTTTCAGTGACAACTCGA 6934

QY 4693 GCACAGTGGCAAGAACGCCCGTCTGTGGCCAGCAAGATAGCGCGTCTCGTCT 4752
Db 6933 GCACAGTGGCAAGAACGCCCGTCTGTGGCCAGCAAGATAGCGCGTCTCGTCT 6874
QY 4753 GCAGTTCAATTCAGGCGACCGGACAGGTCGCTCTTGACAAAAGAACCGGCGCCCTGCG 4812
Db 6873 GCAGTTCAATTCAGGCGACCGGACAGGTCGCTCTTGACAAAAGAACCGGCGCCCTGCG 6814
QY 4813 CTGACAGCCGGAACACGCGCGCATCAGAGCAGCGATGCTGTTGTCGCGCATAGC 4872
Db 6813 CTGACAGCCGGAACACGCGCGCATCAGAGCAGCGATGCTGTTGTCGCGCATAGC 6754
QY 4873 CGATAGCTCTCCACCCAGCGCGGAGAACCTGGTGCAATCCATCTTGTTCATCA 4932
Db 6753 CGATAGCTCTCCACCCAGCGCGGAGAACCTGGTGCAATCCATCTTGTTCATCA 6694
QY 4933 TGGAAACGATCTCATCTCTCTCTTGTATCAGATCTTG 4971
Db 6693 TGGAAACGATCTCATCTCTCTCTTGTATCAGATCTTG 6655

RESULT 13

PCT-US96-10041-6/c
; Sequence 6, Application PC/TUS9610041
; GENERAL INFORMATION:
; APPLICANT: White Phd, Kenneth
; APPLICANT: Morrey Phd, John
; APPLICANT: Reed, William
; TITLE OF INVENTION: Cassette for Expression of Lytic
; TITLE OF INVENTION: Peptides in Mammalian Transgenic Organisms
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Trask Britt and Rossa
; STREET: P.O. Box 2550
; CITY: Salt Lake City
; STATE: Utah
; COUNTRY: USA
; ZIP: 84110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10041
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sweigert Phd, Susan E
; REGISTRATION NUMBER: 36,289
; REFERENCE/DOCKET NUMBER: 2549
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 8015321922
; TELEFAX: 8015319168
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8797 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Construct comprising Bos
; DESCRIPTION: taurus beta casein 5' regulatory region plus genes encoding
; DESCRIPTION: amphipathic peptide and green fluorescent protein"
; HYPOTHEtical: YES
; ANTI-SENSE: NO

PCT-US96-10041-6

Query Match 29.7%; Score 1569.8; DB 5; Length 8797;

Best Local Similarity 86.3%; Pred. No. 0;

Matches 1847; Conservative 0; Mismatches 112; Indels 180; Gaps 3;

QY 3013 MAGCGGTAATACGGTTATCCACAGATCAAGGGGATAACGAGGAAACATGTGAGCA 3072

Db 8793 ATGCGGTAATACGGTATATCAAGATCAGGGATACACAGAAAGAAATGTGAGCA 8734
Qy
Db 3073 AAGGCGAGCAAAAGGCGAGAAACCGTAAAGAGCGCGGTGTGTGCGTTTTCATAGG 3132
Qy
Db 8733 AAGGCGAGCAAAAGGCGAGAAACCGTAAAGAGCGCGGTGTGTGCGTTTTCATAGG 8674
Qy
Db 3133 CTCGCGCCCTGACAGAGCATCACAAAATCGAGCTCAAGTCAGAGGTGGCGAAACCG 3192
Qy
Db 8673 CTCGCGCCCTGACAGAGCATCACAAAATCGAGCTCAAGTCAGAGGTGGCGAAACCG 8614
Qy
Db 3193 ACAGGACTATAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTGTT 3252
Qy
Db 8613 ACAGGACTATAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTGTT 8554
Qy
Db 3253 CCGACCTCGCGTTACCGGATACCTGTCGCGCTTCTCCCTTGGGAGCGTGGCGCTT 3312
Qy
Db 8553 CCGACCTCGCGTTACCGGATACCTGTCGCGCTTCTCCCTTGGGAGCGTGGCGCTT 8494
Qy
Db 3313 TCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGTTGAGTTCGCTCCAAAGCTGGG 3372
Qy
Db 8493 TCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGTTGAGTTCGCTCCAAAGCTGGG 8434
Qy
Db 3373 TGTGTGACGAAACCCCGGTTACGCGGACCGGCTGCGCTTATCCGTAAGTATCTCTT 3432
Qy
Db 8433 TGTGTGACGAAACCCCGGTTACGCGGACCGGCTGCGCTTATCCGTAAGTATCTCTT 8374
Qy
Db 3433 GAGTCCAAACCGGTAAGACAGACTATCGCACTGGCAGCAGCACTGGTAACAGGATT 3492
Qy
Db 8373 GAGTCCAAACCGGTAAGACAGACTATCGCACTGGCAGCAGCACTGGTAACAGGATT 8314
Qy
Db 3493 AGCAGAGGAGGTATGAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCTTAAGTACG 3552
Qy
Db 8313 AGCAGAGGAGGTATGAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCTTAAGTACG 8254
Qy
Db 3553 TACACTAGAAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGATTACTTCGGA 3612
Qy
Db 8253 TACACTAGAAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGATTACTTCGGA 8194
Qy
Db 3613 AGAGTGTAGTCTTTGATTCGCGCAAAACCAACCGTGTGAGCGGTGTTTTTTGTT 3672
Qy
Db 8193 AGAGTGTAGTCTTTGATTCGCGCAAAACCAACCGTGTGAGCGGTGTTTTTTGTT 8134
Qy
Db 3673 TGCAAGCAGAGATTTACCGCAGAAAAAGATCTCAAGAGATCTTTGATCTTTTCT 3732
Qy
Db 8133 TGCAAGCAGAGATTTACCGCAGAAAAAGATCTCAAGAGATCTTTGATCTTTTCT 8074
Qy
Db 3733 ACGGGTCTGACGCTCAGTGAACGAAACCTACGTTAAGGATTTTGGTCAATGAGATTA 3792
Qy
Db 8073 ACGGGTCTGACGCTCAGTGAACGAAACCTACGTTAAGGATTTTGGTCAATGAGATTA 8014
Qy
Db 3793 TCGTGCACCAAGCGGCGATCGTGCCT----- 3819
Db 8013 TCAAAAGGATCTTCACCTAGATCTTTTAAATTAAGTTTAAATCAATCTAA 7954
Qy
Db 3820 -----CCCCACTCTCAGATTTCGGGGGATG 3845
Qy
Db 7953 AGTATATAGTAACCTGAGGCTATGCGAGGCGTGGCGCCGACGTTGGCTGCGAG 7894
Qy
Db 3846 GATGCGCGATAGCGCTGCTGGTTTCTTGATGCGGACGATTTGACTGCGG----- 3900
Qy
Db 7893 CCTGGGCTTACCCGAACTTTGGGGGTGGGGTGGGAAAGGAAACCGGGCGTAT 7834
Qy
Db 3901 -----TAGAACTCGCGAGGTGCTCCAGCTCAGGACGAGCTGAACCAACTCGCG 3951
Qy
Db 7833 TGGCCCAATGGGTCTCGTGGGTATCGACAGAGTCCAGCCCTGGGACCGAACCCTG 7774
Qy
Db 3952 AGGGATCGA----- 3961
Qy
Db 7773 CGTTTATGAACAAAGACCAACACCGTGGTTTTTATCTGTCTTTTATTTGCGGTCA 7714
Qy
Db 3962 -----GCCCGGGTGG 3972

RESULT 14

US-08-723-306-5/c

Db 7113 GCGCGGTTCTTCCGGTATTGTCTCTCGTGTTTTTCAGTTAGCTTCCCTCAGGCTGG 7654
Qy
Db 3973 CCGAAGAACTCCAGATGAGATCCCGCGCTGGAGGATCATCCAGCGGGTCCCGGAAA 4032
Qy
Db 7653 CCGAAGAACTCCAGATGAGATCCCGCGCTGGAGGATCATCCAGCGGGTCCCGGAAA 7594
Qy
Db 4033 ACAGTTCCGAAGCCCAACCTTTTATAGAAGCGCGGTGGAATCGAAATCTCGTGTGGC 4092
Qy
Db 7593 ACAGTTCCGAAGCCCAACCTTTTATAGAAGCGCGGTGGAATCGAAATCTCGTGTGGC 7534
Qy
Db 4093 AGTTTGGGCTCGCTTGGTGGTCAATTTTCGAAACCCAGAGTCCCGCTCAGAAAGTCTGT 4152
Qy
Db 7533 AGTTTGGGCTCGCTTGGTGGTCAATTTTCGAAACCCAGAGTCCCGCTCAGAAAGTCTGT 7474
Qy
Db 4153 CAAGAAGCGGATAGAAGCGGATGCGCTGCGAATCGGAGCGGCGATACCGTAAAGCAGA 4212
Qy
Db 7473 CAAGAAGCGGATAGAAGCGGATGCGCTGCGAATCGGAGCGGCGATACCGTAAAGCAGA 7414
Qy
Db 4213 GGAAGCGGTACGCGATTTGCGCGCAAGCTCTTTCAGCAATATCAGGATGAGTCCAGCGTA 4272
Qy
Db 7413 GGAAGCGGTACGCGATTTGCGCGCAAGCTCTTTCAGCAATATCAGGATGAGTCCAGCGTA 7354
Qy
Db 4273 TGTCTGTATAGCGGTCCGCGCACACCCAGCGCGCCACAGTTCGATGAATCCAGAAAGCGGC 4332
Qy
Db 7353 TGTCTGTATAGCGGTCCGCGCACACCCAGCGCGCCACAGTTCGATGAATCCAGAAAGCGGC 7294
Qy
Db 4333 CATTTTCCACATGATATTCGGCAAGCGGATTCGCCATGCGTTCAGCAAGTATCCAGT 4392
Qy
Db 7293 CATTTTCCACATGATATTCGGCAAGCGGATTCGCCATGCGTTCAGCAAGTATCCAGT 7234
Qy
Db 4393 CGTGGGATGCGCGCTTGGCTGCGCAACAGTTTCGGCTGGCGCGAGCCCTGATGCT 4452
Qy
Db 7233 CGTGGGATGCTCGCTTGGCTGCGCAACAGTTTCGGCTGGCGCGAGCCCTGATGCT 7174
Qy
Db 4453 CTTCTGCAGATCATCTGATCGAACAGCCGGTTTCATCCGAGTACGTCGCTGCTCGA 4512
Qy
Db 7173 CTTCTGCAGATCATCTGATCGAACAGCCGGTTTCATCCGAGTACGTCGCTGCTCGA 7114
Qy
Db 4513 TCGGATGTTTCGCTTGGTGGTTCGATGGCAGGTAGCCGATCAAGCTATGACGCCGCC 4572
Qy
Db 7113 TCGGATGTTTCGCTTGGTGGTTCGATGGCAGGTAGCCGATCAAGCTATGACGCCGCC 7054
Qy
Db 4573 GCATTCATCAGCCATGATGATACTTCTCGGAGGAGCAAGGTGAGTACAGGAGAT 4632
Qy
Db 7053 GCATTCATCAGCCATGATGATACTTCTCGGAGGAGCAAGGTGAGTACAGGAGAT 6994
Qy
Db 4633 CTTGCGCGGCACTTTCGCCATAGCAGCGCTTCCCGCTTCAGTGAACAGCTCGA 4692
Qy
Db 6993 CTTGCGCGGCACTTTCGCCATAGCAGCGCTTCCCGCTTTCAGTGAACAGCTCGA 6934
Qy
Db 4693 GCACAGCTGCGCAAGGAAACCGCGTTCGTCGCGCAGCCACGATAGCGCGCTGCTGCT 4752
Qy
Db 6933 GCACAGCTGCGCAAGGAAACCGCGTTCGTCGCGCAGCCACGATAGCGCGCTGCTGCT 6874
Qy
Db 4753 GCATTCATCAGGCAACCGGACAGTTCGTTGACAAAGAACCGCGCGCCCTCGG 4812
Qy
Db 6873 GCATTCATCAGGCAACCGGACAGTTCGTTGACAAAGAACCGCGCGCCCTCGG 6814
Qy
Db 4813 CTGACAGCGGAAACAGCGCGCATCAGAGCAGCCGATTTGTCTGTTGTGCCAGTCAAGC 4872
Qy
Db 6813 CTGACAGCGGAAACAGCGCGCATCAGAGCAGCCGATTTGTCTGTTGTGCCAGTCAAGC 6754
Qy
Db 4873 CGAATAGCTCTTCACCCAGCGCGGAGAACCTGCGTGCATCATCTTGTTCATCA 4932
Qy
Db 6753 CGAATAGCTCTTCACCCAGCGCGGAGAACCTGCGTGCATCATCTTGTTCATCA 6694
Qy
Db 4933 TCGGAAGCATCTCATCTCTCTTTCATCAGATCTTG 4971
Qy
Db 6693 TCGGAAGCATCTCATCTCTCTTTCATCAGATCTTG 6655

```

; OTHER INFORMATION: /product= "Shiva-1 coding sequence"
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 7630..7635
; OTHER INFORMATION: /standard name= "Bovine beta casein"
; OTHER INFORMATION: 3' region, in exon 9"
US-08-723-306-5

Query Match      29.7%; Score 1569.8; DB 2; Length 11093;
Best Local Similarity 86.3%; Pred. No. 0;
Matches 1847; Conservative 0; Mismatches 112; Indels 180; Gaps 3;

QY 3013 AAGGCGGTAAATACGGTTATCCACAGAAATCAGGGGATTAACGACAGAAAGAAACATCTGAGCA 3072
Db 11089 ATGGCGGTAAATACGGTTATCCACAGATCAGGGGATTAACGCGAAGAAACATCTGAGCA 11030

QY 3073 AAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGGTTGCTGGCGTTTTTCCATAGG 3132
Db 11029 AAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGGTTGCTGGCGTTTTTCCATAGG 10970

QY 3133 CTCGGCCCCCTGACGAGCATCACAAAATCGACGCTCMAGTCAGAGGTGCGGAAACCCG 3192
Db 10969 CTCGGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGCGGAAACCCG 10910

QY 3193 ACAGGACTATAAGATACACAGCGGTTTCCCCCTCGAAAGTCCTCTGTCGCGCTCTCCCTGTT 3252
Db 10909 ACAGGACTATAAGATACACAGCGGTTTCCCCCTCGAAGCTCCCTCGTCGCTCTCTCTGTT 10850

QY 3253 CGGACCCCTGCGCTTACCGGATACCTGTCGCCCTTTCTCCCTTCGGGAAGCGTGCGCGTT 3312
Db 10849 CGGACCCCTGCGCTTACCGGATACCTGTCGCCCTTTCTCCCTTCGGGAAGCGTGCGCGTT 10790

QY 3313 TCTCATAGCTACGCTGTAGGTATCTCAGTTTCGGGTAGGTGTCGTTTCGCTCCAAGCTGGGC 3372
Db 10789 TCTCATAGCTACGCTGTAGGTATCTCAGTTTCGGGTAGGTGTCGTTTCGCTCCAAGCTGGGC 10730

QY 3373 TGTGTGACGAACCCCGCGTTTCAGCCGACCGCTGCGCGTTATCCGGTAACATATCGTCTT 3432
Db 10729 TGTGTGACGAACCCCGCGTTTCAGCCGACCGCTGCGCGTTATCCGGTAACATATCGTCTT 10670

QY 3433 GAGTCCACCCGGTAAGACACGACTTATCGCCACTGCGCAGCAGCCACTGGTAAACAGAGATT 3492
Db 10669 GAGTCCACCCGGTAAGACACGACTTATCGCCACTGCGCAGCAGCCACTGGTAAACAGAGATT 10610

QY 3493 AGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTTGAAGTGGTGCCCTAACTACGGC 3552
Db 10609 AGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTTGAAGTGGTGCCCTAACTACGGC 10550

QY 3553 TACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAA 3612
Db 10549 TACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAA 10490

QY 3613 AGAGTTGGTGTAGCTCTTTGATTCGGGCAAAACAAACACCGCTGGTACGGGTGGTTTTTTGTT 3672
Db 10489 AGAGTTGGTGTAGCTCTTTGATTCGGGCAAAACAAACACCGCTGGTACGGGTGGTTTTTTGTT 10430

QY 3673 TGCAGACGACAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATTCCTTTGATCTTTCT 3732
Db 10429 TGCAGACGACAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATTCCTTTGATCTTTCT 10370

QY 3733 AGGGGCTCTGACGCTCAGTGGAAACGAAACATCAGTTTAAAGGATTTTGGTCATCAGATT 3792
Db 10369 AGGGGCTCTGACGCTCAGTGGAAACGAAACATCAGTTTAAAGGATTTTGGTCATCAGATT 10310

QY 3793 TCGTCGACCAAGACGGCCATCGTGCCT----- 3819
Db 10309 TCAAAAAGGATCTTCCACCTAGATCCTTTTAAATTAATAAATGAAGTTTAAATCAATCTAA 10250

QY 3820 -----CCCCACTCTCGAGTTTCGGGGGCGATG 3845
Db 10249 AGTATATATAGTAACCTGAGGCTATGGCAGGGGCTGCGGCCGACGCTTGGCTGCGAGC 10190

QY 3846 GATGCGGATAGCGGCTGCTGTTTCTGATGTCGACGAGATTGCACTCCCG----- 3900

```

Db 10189 CCTGGGCTTCAACCGAATTTGGGGGTGGGGAAGGAGAAACGCGGGCGTAT 10130
QY 3901 -----TGAAGTCTGGAGAGTGTCTCAGCTCAGGACAGAGTGAACCAACTCGCG 3951
Db 10129 TGGCCCCAATGGGTCTCGTGGGGTATCGACAGAGTGCAGCCCTGGGACGAAACCCG 10070
QY 3952 AGGGGATCGA----- 3961
Db 10069 CGTTATGAACAACAGCCCAACACCGTGGGTTTATCTGTCTTTTATTCGGTCATA 10010
QY 3962 -----GCCCGGGTGG 3972
Db 10009 GCGGGGTTCCTTCGGTATTGTCTCCTTCGGTGTTCAGTTAGCCTCCCCCTAGGGTGG 9950
QY 3973 GCGAAGAACTCAGCATGAGATCCCGCGCTGGAGGATCATCCAGCGCGGTCCCGGAAA 4032
Db 9949 GCGAAGAACTCAGCATGAGATCCCGCGCTGGAGGATCATCCAGCGCGGTCCCGGAAA 9890
QY 4033 ACGATTCCGAAGCCCAACTTTCATAGAAGCGCGGTGGAATCGAAATCTCTGTATGGC 4092
Db 9889 ACGATTCCGAAGCCCAACTTTCATAGAAGCGCGGTGGAATCGAAATCTCTGTATGGC 9830
QY 4093 AGGTTGGCGCTGCTGTGGTGGTCTATTCGAACCCAGAGTCCCGCTCAGAACTCGT 4152
Db 9829 AGGTTGGCGCTGCTGTGGTGGTCTATTCGAACCCAGAGTCCCGCTCAGAACTCGT 9770
QY 4153 CAAGAAGCGGATAGAAGCGGATGGCTGCGAATCGGAGCGGGGATACCGTAAGACACGA 4212
Db 9769 CAAGAAGCGGATAGAAGCGGATGGCTGCGAATCGGAGCGGGGATACCGTAAGACACGA 9710
QY 4213 GGAAGCGGTACGCCATTCGCGCGCAAGCTTTCAGCAATATCAGGGTAGCCACGCTA 4272
Db 9709 GGAAGCGGTACGCCATTCGCGCGCAAGCTTTCAGCAATATCAGGGTAGCCACGCTA 9650
QY 4273 TGTCTGTATAGCGGTCCGCCACACCCAGCGGCGCACAGTGTGATGAATCCAGAAAAGCGC 4332
Db 9649 TGTCTGTATAGCGGTCCGCCACACCCAGCGGCGCACAGTGTGATGAATCCAGAAAAGCGC 9590
QY 4333 CATTTTCCACATGATATTCGGCAAGCGGATGCCATGGGTCAACAGAGATCTTCGC 4392
Db 9589 CATTTTCCACATGATATTCGGCAAGCGGATGCCATGGGTCAACAGAGATCTTCGC 9530
QY 4393 CGTCGGGATCGCGCTTGGCTTGGCAACAGTTCGGCTGGCGGAGCGCCCTGATGCT 4452
Db 9529 CGTCGGGATCGCTGCTTGGCTTGGCAACAGTTCGGCTGGCGGAGCGCCCTGATGCT 9470
QY 4453 CTTCGTCAGATCATCTCTGATCGAAGACCGGCTTCATCCGAGTACGTGCTCGCTCGA 4512
Db 9469 CTTCGTCAGATCATCTCTGATCGAAGACCGGCTTCATCCGAGTACGTGCTCGCTCGA 9410
QY 4513 TCGGATGTTTCGTTGGTGGTTCGAATGGGAGGTAGCGGATCAAGGTATGACGCGCC 4572
Db 9409 TCGGATGTTTCGTTGGTGGTTCGAATGGGAGGTAGCGGATCAAGGTATGACGCGCC 9350
QY 4573 GCATTGCATCGACCATGATGATATCTTTCGCGAGGAGCAAGGTGAGATGACAGGAGAT 4632
Db 9349 GCATTGCATCGACCATGATGATATCTTTCGCGAGGAGCAAGGTGAGATGACAGGAGAT 9290
QY 4633 CTGCGCGCGGACATTCGCGCCATAGCAGCCAGTCCCTTCCCGCTTCAGTGACACGTCGA 4592
Db 9289 CTGCGCGCGGACATTCGCGCCATAGCAGCCAGTCCCTTCCCGCTTCAGTGACACGTCGA 9230
QY 4693 GCACAGCTCGCAGAGAAACGCGCTGCTGGCGCAGCCAGATAGCGCGCTGCTCGCTCCT 4752
Db 9229 GCACAGCTCGCAGAGAAACGCGCTGCTGGCGCAGCCAGATAGCGCGCTGCTCGCTCCT 9170
QY 4753 GCAGTTCAATCAGGCAACGAGAGTGTGCTTGAACAAAGAACCGGGGCGCCCTGCG 4812
Db 9169 GCAGTTCAATCAGGCAACGAGAGTGTGCTTGAACAAAGAACCGGGGCGCCCTGCG 9110
QY 4813 CTGACAGCGGAAACCGGGGATCAGAGAGCGGATGCTGTGTGCGCCAGTCATAGC 4872

Db 9109 CTGACAGCGGAAACACGCGGCATCAGAGCAGCGGATTGTCTGTGTGCCAGTCATAGC 9050
QY 4873 CGAATAGCTCTCCACCCAGCGCGGAGAACTGGTGAATCCATCTTGTTCATCA 4932
Db 9049 CGAATAGCTCTCCACCCAGCGCGGAGAACTGGTGAATCCATCTTGTTCATCA 8990
QY 4933 TCGGAAAGCATCTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4971
Db 8989 TCGGAAAGCATCTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8951

RESULT 15
PCT-US96-10041-5/c
; Sequence 5, Application PC/TUS9610041
; GENERAL INFORMATION:
; APPLICANT: White PhD, Kenneth
; APPLICANT: Morrey PhD, John
; APPLICANT: Reed, William
; TITLE OF INVENTION: Cassette for Expression of Lytic
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Trask Britt and Rossa
; STREET: P.O. Box 2550
; CITY: Salt Lake City
; STATE: Utah
; COUNTRY: USA
; ZIP: 84110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10041
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sweigert PhD, Susan E
; REGISTRATION NUMBER: 36,289
; REFERENCE/DOCKET NUMBER: 2549
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 8015321922
; TELEFAX: 8015319168
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11093 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Construct comprising
; DESCRIPTION: portions of Bos taurus beta casein gene and genes encoding
; DESCRIPTION: amphipathic peptide and green fluorescent protein"
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1801..1834
; OTHER INFORMATION: /product= "beta casein exon 1"
; FEATURE:
; NAME/KEY: exon
; LOCATION: 3780..3832
; OTHER INFORMATION: /product= "beta casein exon 2"
; FEATURE:
; NAME/KEY: TATA signal
; LOCATION: 1766..1773
; FEATURE:
; NAME/KEY: exon
; LOCATION: 4567..4590
; OTHER INFORMATION: /product= "portion of beta casein
; OTHER INFORMATION: exon 3"
; FEATURE:

/ NAME/KEY: 5'UTR
/ LOCATION: 127..1800
/ OTHER INFORMATION: /function= "5' flanking regulatory
/ OTHER INFORMATION: region of bovine beta casein gene"
/ FEATURE:
/ NAME/KEY: mat_peptide
/ LOCATION: 4587..5310
/ OTHER INFORMATION: /product= "Green fluorescent
/ OTHER INFORMATION: protein"
/ FEATURE:
/ NAME/KEY: mat_peptide
/ LOCATION: 5320..5449
/ OTHER INFORMATION: /product= "Shiva-1 coding sequence"
/ FEATURE:
/ NAME/KEY: polyA_signal
/ LOCATION: 7630..7635
/ OTHER INFORMATION: /standard_name= "Bovine beta casein
/ OTHER INFORMATION: 3' region, in exon 9"
PCT-US96-10041-5

Query Match 29.7%; Score 1569.8; DB 5; Length 11093;
Best Local Similarity 86.3%; Pred. No. 0;
Matches 1847; Conservative 0; Mismatches 112; Indels 180; Gaps 3;

QY	3013	AAGCGCGTAAATACGGTTATCCAGAAATCAGGGGATAACGAGGAAAGAAATGTGAGCA	3072
DB	11089	ATGGCGGTAAATACGGTTATCCACAGATCAGGGGTAACGAGGAAAGAAATGTGAGCA	11030
QY	3073	AAAGGCCAGCAAAAGCCAGGAACCGTAAAGAGCGCGTGTCTGCGCTTTTCCATAGG	3132
DB	11029	AAAGGCCAGCAAAAGCCAGGAACCGTAAAGAGCGCGTGTCTGCGCTTTTCCATAGG	10970
QY	3133	CTCCGCCCCCTCAGCAGCATCACAAAATCAGCGTCAAGTCAAGTCAAGTCAAGTCAAGT	3192
DB	10969	CTCCGCCCCCTCAGCAGCATCACAAAATCAGCGTCAAGTCAAGTCAAGTCAAGTCAAGT	10910
QY	3193	ACAGGACTATAAGATPACAGGCGTTTCCCTCGTGAAGTCCCTCGTGGCTCTCTGTT	3252
DB	10909	ACAGGACTATAAGATPACAGGCGTTTCCCTCGTGAAGTCCCTCGTGGCTCTCTGTT	10850
QY	3253	CGACCCCTCGGTTACCGGATACCTGTCGGCTTTTCCCTCGGAGTCCCTCGTGGCTCTCTGTT	3312
DB	10849	CGACCCCTCGGTTACCGGATACCTGTCGGCTTTTCCCTCGGAGTCCCTCGTGGCTCTCTGTT	10790
QY	3313	TCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGTGTCTCGTCCAAAGTGGG	3372
DB	10789	TCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGTGTCTCGTCCAAAGTGGG	10730
QY	3373	TGTGTGCAAGAACCCCTCGTTCAGCCCGACCGTGGCGCTTATCCGGTAACTATCGTCTT	3432
DB	10729	TGTGTGCAAGAACCCCTCGTTCAGCCCGACCGTGGCGCTTATCCGGTAACTATCGTCTT	10670
QY	3433	GAGTCCAAACCGGTAAAGACAGCTTATCGCCACTGGCAGCAGCACTGGTAAACAGGATT	3492
DB	10669	GAGTCCAAACCGGTAAAGACAGCTTATCGCCACTGGCAGCAGCACTGGTAAACAGGATT	10610
QY	3493	AGCAGACGAGGTATGTAGCGGTGTCTACAGAGTCTTGAAGTGGTGGCTTAACCTACGGC	3552
DB	10609	AGCAGACGAGGTATGTAGCGGTGTCTACAGAGTCTTGAAGTGGTGGCTTAACCTACGGC	10550
QY	3553	TACACTAGAAGACAGTATTTGGTATCTCGCTCTGCTGTAAGCCAGTTACCTTCGGAAAA	3612
DB	10549	TACACTAGAAGACAGTATTTGGTATCTCGCTCTGCTGTAAGCCAGTTACCTTCGGAAAA	10490
QY	3613	AGAGTTGGTAGTCTTTGATCCCGCAAAACAAACACCGTGGTGGTGGTGGTGGTGGTGGT	3672
DB	10489	AGAGTTGGTAGTCTTTGATCCCGCAAAACAAACACCGTGGTGGTGGTGGTGGTGGTGGT	10430
QY	3673	TGCAAGCAGCAGATTACGGCGAGAAAAAGGATCTCAAGAGATCTTTGATCTTTCT	3732
DB	10429	TGCAAGCAGCAGATTACGGCGAGAAAAAGGATCTCAAGAGATCTTTGATCTTTCT	10370
QY	3733	ACGGGCTCTGACGCTCAGTGTGAACGAAACTCAGGTTAAGGGATTTTGGTCAATGAGATTA	3792

DB	10369	ACGGGCTCTGACGCTCAGTGTGAACGAAACTCAGTTAAGGGATTTTGGTCAATGAGATTA	10310
QY	3793	TGCTCGACCAAGCGGCCATCGTCCCT	3819
DB	10309	TCAAAAAGGATCTTACCTAGATCTCTTTTAAATTAATAAGAGTTTAAATCAATCTAA	10250
QY	3820	-----CCCCACTCTCTGAGTTTCGGAGTTTGCACCTGCCG-----	3845
DB	10249	AGTATATATAGTAACTGAGGCTATGCGAGGCGCTGCGCCCGACGTTGGTGGCAGC	10190
QY	3846	GATCGCGGATAGCGCTGCTGCTTCTTGGATCCGCGGATTTGCACCTGCCG-----	3900
DB	10189	CCTGGGCTTTCACCCGAACCTTGGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGT	10130
QY	3901	-----TAGAACTCCGCGAGTCTCCAGCCTCAGGCGAGCTGAACCAACTCGCG	3951
DB	10129	TGSCCCCAATGGGCTCTCGGTGGGTATCGACAGAGTGCAGCCCTGGGACCGAACCCG	10070
QY	3952	AGGGATCGA-----	3961
DB	10069	CGTTTATGAACAAACGACCAACACCGTGCCTTTTATTCTGCTTTTATTGCGCTATA	10010
QY	3962	-----GCCCGGGTGG	3972
DB	10009	GCGGGGTTCCTTCGGTATTGTCTCTTCCGTGTTTTCAGTTAGCTTCCCTTAGGGTGG	9950
QY	3973	GCAGAAACTCCAGCATGATATCCCGCGCTGGAGATCATCAGCGGGCGTCCCGGAA	4032
DB	9949	GCAGAAACTCCAGCATGATATCCCGCGCTGGAGATCATCAGCGGGCGTCCCGGAA	9990
QY	4033	ACGATTCGGAAGCCCAACCTTTCATAGAGCGCGGTGGAATCGAAATCTGCTGATGGC	4092
DB	9889	ACGATTCGGAAGCCCAACCTTTCATAGAGCGCGGTGGAATCGAAATCTGCTGATGGC	9830
QY	4093	AGGTGGCGCTCGCTTGGTGGTCAATTCGAAACCCAGAGTCCCGTCAAGAACTCGT	4152
DB	9829	AGGTGGCGCTCGCTTGGTGGTCAATTCGAAACCCAGAGTCCCGTCAAGAACTCGT	9770
QY	4153	CAAGAGCGCATAGAGCGATCGCTCGMATCGGAGGGGCGATACCGTAAAGCAGCA	4212
DB	9769	CAAGAGCGCATAGAGCGATCGCTCGMATCGGAGGGGCGATACCGTAAAGCAGCA	9710
QY	4213	GGAGCGGTCAAGCCATTCGCCCGCAAGCTCTTCAGCAATATCACGGGTAGCCAAAGCTA	4272
DB	9709	GGAGCGGTCAAGCCATTCGCCCGCAAGCTCTTCAGCAATATCACGGGTAGCCAAAGCTA	9650
QY	4273	TGCTCTGATAGCGGTCCGCCACACCGAGCGGCGCACTGCTGATGAATCCAGAAAGCGGC	4332
DB	9649	TGCTCTGATAGCGGTCCGCCACACCGAGCGGCGCACTGCTGATGAATCCAGAAAGCGGC	9590
QY	4333	CATTTTCCACCATGATATTCGGCAAGCAGGATCGCCATGGGTACAGAGATCTCTCGC	4392
DB	9589	CATTTTCCACCATGATATTCGGCAAGCAGGATCGCCATGGGTACAGAGATCTCTCGC	9530
QY	4393	CGTGGGCACTGCGCGCTTGAAGCTGGCGAAACAGATTTCGCTGGCGAGCCCTGATGCT	4452
DB	9529	CGTGGGCACTGCGCGCTTGAAGCTGGCGAAACAGATTTCGCTGGCGAGCCCTGATGCT	9470
QY	4453	CTTCTGCGAGATCATCTGATCGAAGACCGGCTTCATCCGAGTACGCTCGCTCGCA	4512
DB	9469	CTTCTGCGAGATCATCTGATCGAAGACCGGCTTCATCCGAGTACGCTCGCTCGCA	9410
QY	4513	TGCGATGTTTTGCTGGTGGTGGAAATGGGCGAGTAGCCGATCAAGCGTATGCAGCGGC	4572
DB	9409	TGCGATGTTTTGCTGGTGGTGGAAATGGGCGAGTAGCCGATCAAGCGTATGCAGCGGC	9350
QY	4573	GCATTGCATCAGCATGATGATCTTTCTCGCAGAGAGAGGTTGAGTACAGGAGAT	4632
DB	9349	GCATTGCATCAGCATGATGATCTTTCTCGCAGAGAGAGGTTGAGTACAGGAGAT	9290
QY	4633	CTGCCCCCGCACTTCCGCCAATAGCAGCGAGTCCCTTCCGCTTCAGTGAACAGCTCGA	4692

Db 9289 CCTGCCCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCCCGCTTCAGTGACAAACGTGGA 9230
Qy 4893 GCACAGCTGCCAAGGAACGCCCGCTCGTGGCCACGATAGCCGGCGCTGCGCTCGTCT 4752
Db 9229 GCACAGCTGCCAAGGAACGCCCGCTCGTGGCCACGATAGCCGGCGCTGCGCTCGTCT 9170
Qy 4753 GCAGTTAATTCAGGGCACCGGACAGGTGCGTCTTGACAAAAGAACCGGGCGGCCCTGGG 4812
Db 9169 GCAGTTCAATTCAGGGCACCGGACAGGTGCGTCTTGACAAAAGAACCGGGCGGCCCTGGG 9110
Qy 4813 CTGACAGCCGGAACACGCGGCGCATCAGAGCAGCCGATTGTCTGTGTGCCCGAGTCATAGC 4872
Db 9109 CTGACAGCCGGAACACGCGGCGCATCAGAGCAGCCGATTGTCTGTGTGCCCGAGTCATAGC 9050
Qy 4873 CGAATAGCTCTCCACCCAGCCGCGGAGAACCTGGTGCAATCCATCTTGTTCATCA 4932
Db 9049 CGAATAGCTCTCCACCCAGCCGCGGAGAACCTGGTGCAATCCATCTTGTTCATCA 8990
Qy 4933 TGCAGAACGATCCTCATCCTGCTCTTGTGATCAGATCTTG 4971
Db 8989 TGCAGAACGATCCTCATCCTGCTCTTGTGATCAGATCTTG 8951

Search completed: February 15, 2004, 05:49:54
Job time : 350 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 05:29:07 ; Search time 1682 Seconds
(without alignments)
11569.936 Million cell updates/sec

Title: US-09-921-143-36
Perfect score: 5283
Sequence: 1 agcttgacctatgcgact.....tgatgcttgaggagcgtg 5283

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2449703 seqs, 1841816367 residues

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubna/US05_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq:
14: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq:
17: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq:
18: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5283	100.0	5283	13	US-09-921-143-36
2	2240	42.4	8349	13	US-10-198-478-16
3	1793.4	33.9	4800	13	US-10-322-360-1
4	1692.4	32.0	5225	13	US-10-286-186-16
5	1632.8	30.9	5285	14	US-10-038-001-1
6	1603.8	30.4	2192	15	US-10-021-403A-10
7	1603.8	30.4	3534	12	US-10-395-709-11
8	1603.8	30.4	3534	12	US-10-395-709-12
9	1603.8	30.4	3534	12	US-10-395-709-13
10	1603.8	30.4	3534	12	US-10-395-709-14
11	1603.8	30.4	3534	12	US-10-395-709-15
12	1603.8	30.4	3534	15	US-10-021-403A-9
13	1600	30.3	5966	13	US-10-136-837-3
14	1571.6	29.7	4665	9	US-09-759-960-7
15	1570	29.7	4886	15	US-10-128-853-4

C 16	1570	29.7	5031	15	US-10-033-717-9	Sequence 9, Appl
C 17	1570	29.7	5054	15	US-10-033-717-8	Sequence 8, Appl
C 18	1570	29.7	5130	15	US-10-033-717-7	Sequence 7, Appl
C 19	1570	29.7	5168	15	US-10-033-717-5	Sequence 5, Appl
C 20	1569.8	29.7	4692	15	US-10-161-403-29	Sequence 29, Appl
C 21	1569.8	29.7	4727	13	US-10-115-987B-12	Sequence 12, Appl
C 22	1569.8	29.7	4733	11	US-09-797-496B-1	Sequence 1, Appl
C 23	1569.8	29.7	4733	13	US-10-177-390-1	Sequence 1, Appl
C 24	1569.8	29.7	4740	13	US-10-115-987B-13	Sequence 13, Appl
C 25	1569.8	29.7	4748	13	US-09-796-575-4	Sequence 4, Appl
C 26	1569.8	29.7	4862	15	US-10-161-403-87	Sequence 87, Appl
C 27	1569.8	29.7	4894	15	US-10-055-794-2	Sequence 2, Appl
C 28	1569.8	29.7	4992	13	US-09-796-575-5	Sequence 5, Appl
C 29	1569.8	29.7	5166	15	US-10-033-717-6	Sequence 6, Appl
C 30	1569.8	29.7	5953	12	US-10-157-317-13	Sequence 13, Appl
C 31	1569.8	29.7	5953	12	US-10-157-317-13	Sequence 13, Appl
C 32	1569.8	29.7	5953	13	US-10-157-305A-13	Sequence 13, Appl
C 33	1569.8	29.7	5953	13	US-10-157-391-13	Sequence 13, Appl
C 34	1569.8	29.7	5953	13	US-10-157-095-13	Sequence 13, Appl
C 35	1569.8	29.7	5953	13	US-10-157-215A-13	Sequence 13, Appl
C 36	1569.8	29.7	5953	13	US-10-157-309-13	Sequence 13, Appl
C 37	1569.8	29.7	5953	13	US-10-154-951B-13	Sequence 13, Appl
C 38	1569.8	29.7	5953	13	US-10-157-299-13	Sequence 13, Appl
C 39	1569.8	29.7	5953	13	US-10-156-831-13	Sequence 13, Appl
C 40	1569.8	29.7	5953	13	US-10-156-902-13	Sequence 13, Appl
C 41	1569.8	29.7	5953	13	US-10-157-147-13	Sequence 13, Appl
C 42	1569.8	29.7	5953	13	US-10-157-166-13	Sequence 13, Appl
C 43	1569.8	29.7	5953	13	US-10-157-318-13	Sequence 13, Appl
C 44	1569.8	29.7	5953	13	US-10-156-811-13	Sequence 13, Appl
C 45	1569.8	29.7	5953	13	US-10-157-320A-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-09-921-143-36
; Sequence 36, Application US/09921143
; Publication No. US20030215921A1
; GENERAL INFORMATION:
; APPLICANT: Coleman, Timothy
; TITLE OF INVENTION: Vascular Endothelial Growth Factor-2
; FILE REFERENCE: PFI12P6
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/223,276
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 36
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-921-143-36

Query Match 100.0%; Score 5283; DB 13; Length 5283;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ARCTTGACCTTATGGACCTTTCCTACTTGGCAGTACATCTAGTATTAGTCATCGCTAT	60
Db	1	ARCTTGACCTTATGGACCTTTCCTACTTGGCAGTACATCTAGTATTAGTCATCGCTAT	60
Qy	61	TACCATGGTGGTATCGGCTTTTGGCAGTACATCAATGGCGGTGGATCGGTTTGGTACACG	120
Db	61	TACCATGGTGGTATCGGCTTTTGGCAGTACATCAATGGCGGTGGATCGGTTTGGTACACG	120
Qy	121	GGGATTTCCAAAGTCTCCACCTGAGTCAATGGAGTTTGTGTCACCAAAATCA	180
Db	121	GGGATTTCCAAAGTCTCCACCTGAGTCAATGGAGTTTGTGTCACCAAAATCA	180
Qy	181	ACGAGACTTTCCAAATGTCGTAAACATCCCGCCCATTTGACGCAATGGCGGTGGTACG	240

QY 2401 GCAGTGGCAACAAGTGGGTGGGTGGAGCCCGGGGCGAGGTGACCTTCAGACCTTGGCA 2460
Db 2401 GCAGTGGCAACAAGTGGGTGGGTGGAGCCCGGGGCGAGGTGACCTTCAGACCTTGGCA 2460
QY 2461 CTGGAGGTGGCCGGGCGAGAGCGCGGCATCGTGGATCAGTGTGACACAGCATCTGCTCT 2520
Db 2461 CTGGAGGTGGCCGGGCGAGAGCGCGGCATCGTGGATCAGTGTGACACAGCATCTGCTCT 2520
QY 2521 CTCTACCAACTGGAGAACCTTACGCAACTAGGCCACCACTACCTGTGCAACCCCTCTGCA 2580
Db 2521 CTCTACCAACTGGAGAACCTTACGCAACTAGGCCACCACTACCTGTGCAACCCCTCTGCA 2580
QY 2581 ATGAATAAACCCTTTGAAAGAGCACTACAAGTTGTGTGTACATGCTGTGATGTCATATG 2640
Db 2581 ATGAATAAACCCTTTGAAAGAGCACTACAAGTTGTGTGTACATGCTGTGATGTCATATG 2640
QY 2641 TGGTGGGGGGGAAACATGATGGGGTGGCTGGAGTGTGCGGGCTTAATCTATCTGGCA 2700
Db 2641 TGGTGGGGGGGAAACATGATGGGGTGGCTGGAGTGTGCGGGCTTAATCTATCTGGCA 2700
QY 2701 GCTGTCTAGACGTAAATCATGCTCATAGCTGTTTCTGTGTGAATGTTATCCGCTCACA 2760
Db 2701 GCTGTCTAGACGTAAATCATGCTCATAGCTGTTTCTGTGTGAATGTTATCCGCTCACA 2760
QY 2761 ATTCCACACAACATAGAGCGGAAGCATAAAGTCTAAAGCTGGGGTGCCTAATGAGTG 2820
Db 2761 ATTCCACACAACATAGAGCGGAAGCATAAAGTCTAAAGCTGGGGTGCCTAATGAGTG 2820
QY 2821 AGCTAACTACATAATTAATTCGTTGGCTCAGTCCCGCTTTCCAGTCGGGAACCTGTGCG 2880
Db 2821 AGCTAACTACATAATTAATTCGTTGGCTCAGTCCCGCTTTCCAGTCGGGAACCTGTGCG 2880
QY 2881 TGCCAGCTGCAATTAATGAATGCGCAACGCGGGGAGAGCGGTTTCGCTAATTTGGGCGC 2940
Db 2881 TGCCAGCTGCAATTAATGAATGCGCAACGCGGGGAGAGCGGTTTCGCTAATTTGGGCGC 2940
QY 2941 TCTTTCGCTTCTCCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000
Db 2941 TCTTTCGCTTCTCCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000
QY 3001 TCAGCTCACTCAAGCGGTAAATACGTTATCCACAGAAATCAGGGGATAACGCGAGGAAG 3060
Db 3001 TCAGCTCACTCAAGCGGTAAATACGTTATCCACAGAAATCAGGGGATAACGCGAGGAAG 3060
QY 3061 AACATGTGAGCAAAAGCCAGCAAGCCAGCAAGCCAGCAAGCCAGCAAGCCAGCAAGCCAG 3120
Db 3061 AACATGTGAGCAAAAGCCAGCAAGCCAGCAAGCCAGCAAGCCAGCAAGCCAGCAAGCCAG 3120
QY 3121 TTTTTCCTAGCTCGGCCCTGACGAGCATCAGGATCAGAAATCAGCGTCAAGTCAGAGG 3180
Db 3121 TTTTTCCTAGCTCGGCCCTGACGAGCATCAGGATCAGAAATCAGCGTCAAGTCAGAGG 3180
QY 3181 TGGCGAAACCCGACAGGACTATAAGATACAGGCGTTTCCCTCGAAGCTCCCTCGTG 3240
Db 3181 TGGCGAAACCCGACAGGACTATAAGATACAGGCGTTTCCCTCGAAGCTCCCTCGTG 3240
QY 3241 CGCTCTCCTGTTCCGACCGCTGACGAGTACCGGATCAGGATCAGGATCAGGATCAGGAT 3300
Db 3241 CGCTCTCCTGTTCCGACCGCTGACGAGTACCGGATCAGGATCAGGATCAGGATCAGGAT 3300
QY 3301 AGCGTGGCGCTTCTCATAGCTCAGCTGTAGTATCTCAGTTCGGGTAGGTCGTTGCGC 3360
Db 3301 AGCGTGGCGCTTCTCATAGCTCAGCTGTAGTATCTCAGTTCGGGTAGGTCGTTGCGC 3360
QY 3361 TCCAAAGTGGGTGTGTGCAAGAACCCCGCTCAGCCCGAGCGCTGCGCTTATCCGGT 3420
Db 3361 TCCAAAGTGGGTGTGTGCAAGAACCCCGCTCAGCCCGAGCGCTGCGCTTATCCGGT 3420
QY 3421 AACTATCTCTTGTAGTCCAAACCGGTAGACAGCACTTATCCGACCTGGAGAGCCACT 3480
Db 3421 AACTATCTCTTGTAGTCCAAACCGGTAGACAGCACTTATCCGACCTGGAGAGCCACT 3480
QY 3481 GGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTTCTTGAAGTGGTG 3540

Db 3481 GGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTTCTTGAAGTGGTG 3540
QY 3541 CCTAACTACGGCTACACTAGAGAAACAGATATTTGGTATCTGCGCTCTGCTGAAAGCCAGTT 3600
Db 3541 CCTAACTACGGCTACACTAGAGAAACAGATATTTGGTATCTGCGCTCTGCTGAAAGCCAGTT 3600
QY 3601 ACCTTCGGAAGAAAGTGTGTAGCTCTTGTATCCGGCAAAACAAACACCGCTGTAGCGGT 3660
Db 3601 ACCTTCGGAAGAAAGTGTGTAGCTCTTGTATCCGGCAAAACAAACACCGCTGTAGCGGT 3660
QY 3661 GGTTTTTTCTTTTGAAGCAGCAGATACGCGCAGAAAAAAGATCTCAAGAGATCCT 3720
Db 3661 GGTTTTTTCTTTTGAAGCAGCAGATACGCGCAGAAAAAAGATCTCAAGAGATCCT 3720
QY 3721 TTGATCTTTTCTACCGGGTCTGACGCTCAGTGTGAACGAAACACTCAAGTTAAAGGATTTTG 3780
Db 3721 TTGATCTTTTCTACCGGGTCTGACGCTCAGTGTGAACGAAACACTCAAGTTAAAGGATTTTG 3780
QY 3781 GTCATGAGATATCTGACCAAGCGGCATCTGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 3840
Db 3781 GTCATGAGATATCTGACCAAGCGGCATCTGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 3840
QY 3841 GCATGATGCGCGATAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3900
Db 3841 GCATGATGCGCGATAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3900
QY 3901 TAGAATCTCCGCGAGGTGCTCCAGCTCAGCAGCAGCTGAAACCACTCCGCGAGGGATCG 3960
Db 3901 TAGAATCTCCGCGAGGTGCTCCAGCTCAGCAGCAGCTGAAACCACTCCGCGAGGGATCG 3960
QY 3961 AGCCCGGGTGGCGAAGAACTCCAGCATGAGATCCCGCTGCGAGGATCATCCAGCGG 4020
Db 3961 AGCCCGGGTGGCGAAGAACTCCAGCATGAGATCCCGCTGCGAGGATCATCCAGCGG 4020
QY 4021 GCTCTCCGGAACAGATTCCGAAGCCAACTTTTATAGAAAGCGCGGTGGAATCGAAA 4080
Db 4021 GCTCTCCGGAACAGATTCCGAAGCCAACTTTTATAGAAAGCGCGGTGGAATCGAAA 4080
QY 4081 TCTCTGTGAGGCGAGGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4140
Db 4081 TCTCTGTGAGGCGAGGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4140
QY 4141 AGAAGAACTCTGCAAGAGCGGATAGAGCGCATGCTGCGAATCGGAGCGGGCATAC 4200
Db 4141 AGAAGAACTCTGCAAGAGCGGATAGAGCGCATGCTGCGAATCGGAGCGGGCATAC 4200
QY 4201 CGTAAAGCAGAGAGAGCGGTGAGCCATTCGCCCGCAAGCTCTTCAGCAATATCACGGG 4260
Db 4201 CGTAAAGCAGAGAGAGCGGTGAGCCATTCGCCCGCAAGCTCTTCAGCAATATCACGGG 4260
QY 4261 TAGCCAACTGCTATGCTGATAGCGTCCGCGCACACCCGCGGCGCACAGTCGATGAATC 4320
Db 4261 TAGCCAACTGCTATGCTGATAGCGTCCGCGCACACCCGCGGCGCACAGTCGATGAATC 4320
QY 4321 CAGAAAGCGGCCATTTTCCACCATGATATTCGCAAGCAGGCATCGCCATGGGTCAAGA 4380
Db 4321 CAGAAAGCGGCCATTTTCCACCATGATATTCGCAAGCAGGCATCGCCATGGGTCAAGA 4380
QY 4381 CGAGATCTCTCGCGCTGGGCGATCGCGCTGAGCTGCGCAACAGTTCCGGTGGGCGGA 4440
Db 4381 CGAGATCTCTCGCGCTGGGCGATCGCGCTGAGCTGCGCAACAGTTCCGGTGGGCGGA 4440
QY 4441 GCCCTGTAGTCTTCTGCTCCAGATCATCTGATCGCAAGACCGGCTTCATCCGAGTAC 4500
Db 4441 GCCCTGTAGTCTTCTGCTCCAGATCATCTGATCGCAAGACCGGCTTCATCCGAGTAC 4500
QY 4501 GTGCTCGCTCGATCGATGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4560
Db 4501 GTGCTCGCTCGATCGATGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4560
QY 4561 TATGACAGCGCCGCTTTCATCAGCAGCATGATGATCTTCTTCGCGCAGGACAGGTCAG 4620

[illegible]

RESULT 2
 US-10-198-478-16
 ; Sequence 16, Application US/10198478
 ; Publication No. US20030188336A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Corbin, David R.
 ; APPLICANT: Romano, Charles P.
 ; TITLE OF INVENTION: Improved Methods for Transforming Plants to Express delta-Endotoxin
 ; FILE REFERENCE: 38-21 (13547) B
 ; CURRENT APPLICATION NUMBER: US/10/198,478
 ; CURRENT FILING DATE: 2002-11-12
 ; PRIOR APPLICATION NUMBER: 05/186, 002
 ; PRIOR FILING DATE: 1998-11-04
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 16
 ; LENGTH: 8349
 ; TYPE: DNA
 ; ORGANISM: artificial sequence
 ; FEATURE:
 ; OTHER INFORMATION: completely synthesized
 ; FEATURE:

4680	DB	CTGTTCGGACCTCGCGCTTACCGGANATCCTGTCGCGCTTCTCCCTTCGGGAACGGTG	4621
3307	QY	GCCTTTTCTCATAGCTCACGCTCTAGGTATCTCAGTTTCGGTGTAGTCTGTTCCGCTCCAAAG	3366
4620	DB	GCCTTTTCTCATAGCTCACGCTCTAGGTATCTCAGTTTCGGTGTAGTCTGTTCCGCTCCAAAG	4561
3367	QY	CTGGGCTGTGTGCACGAACCCCGGTTACGCCGACGGCTGGCGCTTATCCGGTACTAT	3426
4560	DB	CTGGGCTGTGTGCACGAACCCCGGTTACGCCGACGGCTGGCGCTTATCCGGTAACTAT	4501
3427	QY	CGTCTTTAGTCCAAACCCGGTAAACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC	3486
4500	DB	CGTCTTTAGTCCAAACCCGGTAAACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC	4441
3487	QY	AGGATTAGACAGAGGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGTGGCCTAAC	3546
4440	DB	AGGATTAGACAGAGGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGTGGCCTAAC	4381
3547	QY	TACGGCTACACTAGAAGAACAGTATTGGTGTATCTGGCTCTGCTGAAGCCAGTTACTTTC	3606
4380	DB	TACGGCTACACTAGAAGAACAGTATTGGTGTATCTGGCTCTGCTGAAGCCAGTTACTTTC	4321
3607	QY	GGAAAAGAGTTGGTAGCTCTTGATCGGCAAAACAAACCAACCGCTGGTAGCGGTGTTTT	3666
4320	DB	GGAAAAGAGTTGGTAGCTCTTGATCGGCAAAACAAACCAACCGCTGGTAGCGGTGTTTT	4261
3667	QY	TTTGTGTGAAGCAGCAGATATAGCCGACAGAAAAGGATCTCAAGAGATCCTTTGATC	3726
4260	DB	TTTGTGTGAAGCAGCAGATATAGCCGACAGAAAAGGATCTCAAGAGATCCTTTGATC	4201
3727	QY	TTTTTCTACGGGGTCTGACGCTCAGTGAACGAAAACTCAGTTAAGGGATTTTGGTCATG	3786
4200	DB	TTTTTCTACGGGGTCTGACGCTCAGTGAACGAAAACTCAGTTAAGGGATTTTGGTCAT-	4142
3787	QY	AGATTATCGTCGACCAAGCGGCATCGTGCCTCCCACTCTCTGCAGTTCCGGGGCATGG	3846
4141	DB	-----	4142
3847	QY	ATCGCGGATAGCGCTGCTGTTCTCTGGATGCCGACGGATTGCACTGCCGGTAGAAC	3906
4141	DB	-----	4142
3907	QY	TCCGGAGGTCTGTCCAGCTTCAGGACGACTGAACCAATCCGGAGGGATTCGAGCCCG	3966
4141	DB	-----GCCCG 4137	4137
3967	QY	GGTGGCGGAAGAACTCCAGCATGAGATCCCCGCTGGAGGATCATCCAGCCGGCTCC	4026
4136	DB	GGTGGCGGAAGAACTCCAGCATGAGATCCCCGCTGGAGGATCATCCAGCCGGCTCC	4077
4027	QY	CGGAAAAAGATTCCGAAGCCCAACCTTTATAGAAGCGCGCGTGGAAATCGAAATCTCGT	4086
4076	DB	CGGAAAAAGATTCCGAAGCCCAACCTTTATAGAAGCGCGCGTGGAAATCGAAATCTCGT	4017
4087	QY	GATGGCAGGTTGGCGTTCGTTGGTTCATTTCGAACCCAGAGTCCCGCTCAGAGA	4146
4016	DB	GATGGCAGGTTGGCGTTCGTTGGTTCATTTCGAACCCAGAGTCCCGCTCAGAGA	3957
4147	QY	ACTCGTCAAGAAGCGCATAGAAGCGCATGGCTGCGAATCCGGAGCGCGCATACCGTAAA	4206
3956	DB	ACTCGTCAAGAAGCGCATAGAAGCGCATGGCTGCGAATCCGGAGCGCGCATACCGTAAA	3897
4207	QY	GCACGGAAGCGGTGAGCCGATTCGCGCGCAGACTCTTCAGCAATATCACGGGTAGCCA	4266
3896	DB	GCACGGAAGCGGTGAGCCGATTCGCGCGCAGACTCTTCAGCAATATCACGGGTAGCCA	3837
4267	QY	ACGTATGTCTCTGATAGCGTCCGCCACACCCAGCCGGCCACAGTTCGATGAATCCAGAAA	4326
3836	DB	ACGTATGTCTCTGATAGCGTCCGCCACACCCAGCCGGCCACAGTTCGATGAATCCAGAAA	3777
4327	QY	AGCGGCCATTTTCCACCATGATATTCGGCAAGCAGGCATCCGCATGGGTCAACGACGAT	4386
3776	DB	AGCGGCCATTTTCCACCATGATATTCGGCAAGCAGGCATCCGCATGGGTCAACGACGAT	3717

RESULT 4

RESULT 4
US-10-286-186-16/c

US-10-286-186-16/C
: Sequence 16, Application US/10286186

;; sequence 16, Application US/1028
; Publication No. US20030143741A1

; PUBLICATION NO. 0320
; GENERAL INFORMATION:

APPLICANT: PALMER, KENNETH E.

; AFFILIANT: FARMER, KENNETH
; APPLICANT: POGUE, GARY P.

APPLICANT: (FOOI) GREAT I
; TITLE OF INVENTION: ROLLING CIRCLE REPLICON EXPRESSION

;
; TITLE OF INVENTION: ROCKETING
; TITLE OF INVENTION: VECTOR

FILE REFERENCE: 00801017900US00

FILE REFERENCE: 000101700000
; CURRENT APPLICATION NUMBER: US/10/286,186


```
Db 4145 GGGTCAGTCGTCCGACACACACCGCGGAGTCTCTCCACGAAAGTCCCGGAGAAC 4086
Qy 3983 CCAGCATAGATCCCGCGTGGAGGATCATCCAGCGGGTCCCGGAAACGATTCGA 4042
Db 4085 CGAGCGGTGGTCCAGAACTCGACCGCTCCGCGAGCTCGCGCGGTGAGACCGGAA 4026
Qy 4043 AGCCCAACCTTTCATAGAGCGCGGTGGAATCGAAATCTCGTATGGCAGGTGGGG 4102
Db 4025 CGGCACTGGTCAACTTGGCGATGTGGCCCTCTCTCAGTCTATTATTGAAGCATTTATC 3966
Qy 4103 TCGTTGGTGGTCAATTTGGAACCCAGAGTCCG 4137
Db 3965 AGGCTTATTGTCTCATGAGCGGATACATATTGGAATGTTAGAAAAATAAACAAATAG 3906
Qy 4138 ----- 4137
Db 3905 GGGTCCCGGCACATTTCCCGGAAAGTGCCACTGTATGCGGTGTGAATACCGCACAG 3846
Qy 4138 -----CTCAGAGAACTC 4150
Db 3845 ATGCGTAAGGAGAAAAATACCGCATCAGGAAATTTGAAGCGTTAATAATTCAGAGAACTC 3786
Qy 4151 GTCAGAGCGGATAGAGGATGCGCTGCGATCGGAGCGGCGATACCGTAAAGCAC 4210
Db 3785 GTCAGAGCGGATAGAGGATGCGCTGCGATCGGAGCGGCGATACCGTAAAGCAC 3726
Qy 4211 GAGGAAGCGGTGAGCCCATTCGCGCAAGCTTTTCAGCAATATPCAGGGTAGCCAAAGC 4270
Db 3725 GAGGAAGCGGTGAGCCCATTCGCGCAAGCTTTTCAGCAATATPCAGGGTAGCCAAAGC 3666
Qy 4271 TATGTCTGTAGCGGTCCGCAACCCAGCGGCGACAGTCTGATGTAATCCAGAAAGCG 4330
Db 3665 TATGTCTGTAGCGGTCCGCAACCCAGCGGCGACAGTCTGATGTAATCCAGAAAGCG 3606
Qy 4331 GCCATTTTCCACCATGATATTCGCAAGCAGGCGATCGCATGGTTCACGAGATCCTC 4390
Db 3605 GCCATTTTCCACCATGATATTCGCAAGCAGGCGATCGCATGGTTCACGAGATCCTC 3546
Qy 4391 GCCGTGCGGCGATGCGCGCTTGAAGCTTGGCGAAACAGTTTGGCTGGCGGAGCCCTGATG 4450
Db 3545 GCCGTGCGGCGATGCGCGCTTGAAGCTTGGCGAAACAGTTTGGCTGGCGGAGCCCTGATG 3486
Qy 4451 CTCCTTCGTCCAGATCATCTGATCCAGAACCGCGCTTCCATCCGAGTACGTGCTGCTC 4510
Db 3485 CTCCTTCGTCCAGATCATCTGATCCAGAACCGCGCTTCCATCCGAGTACGTGCTGCTC 3426
Qy 4511 GATGCGATGTTTTCGTTGGTGGTCCAAATGGGAGGATAGCGGATCCGCGATGATGAGCGG 4570
Db 3425 GATGCGATGTTTTCGTTGGTGGTCCAAATGGGAGGATAGCGGATCCGCGATGATGAGCGG 3366
Qy 4571 CCGCATTCGATCAGCATGATGATCTTCTCGCAGGAGCAAGGTGAGATGACAGGAG 4630
Db 3365 CCGCATTCGATCAGCATGATGATCTTCTCGCAGGAGCAAGGTGAGATGACAGGAG 3306
Qy 4631 ATCTGCCCCCGGCACTTCCGCAATAGCAGCGAGTCCCTTCCGCTTCAGTGACAAAGTC 4690
Db 3305 ATCTGCCCCCGGCACTTCCGCAATAGCAGCGAGTCCCTTCCGCTTCAGTGACAAAGTC 3246
Qy 4691 GAGCAGCTGCGCAGGAAACCGCGCTGTCGCGAGCCACGATAGCGGCTGCTGCTC 4750
Db 3245 GAGCAGCTGCGCAGGAAACCGCGCTGTCGCGAGCCACGATAGCGGCTGCTGCTC 3186
Qy 4751 CTCAGTTTCATTCAGGGCACCGGACAGGTGCGTCTTTGACAAAAGAACCGGGCGCCCTG 4810
Db 3185 TTGCACTTCATTCAGGGCACCGGACAGGTGCGTCTTTGACAAAAGAACCGGGCGCCCTG 3126
Qy 4811 CGTTCAGCGCGGAAACAGCGCGGATCAGAGAGCGGATGTCGTTGTCAGGATCAT 4870
Db 3125 CGTTCAGCGCGGAAACAGCGCGGATCAGAGAGCGGATGTCGTTGTCAGGATCAT 3066
Qy 4871 GCCGAATAGCTCTCCACCCAGCGGCGGAGAACTGCGTGTGCAATTCATCTGTTCAAT 4930
```

```
Db 3065 GCCGAATAGCTCTCCACCCAGCGGCGGAGAACCTTGGTGCATCATCTTGTTCAT 3006
Qy 4931 CATGCGAAACGATCCTCATCTCTGTCTTGTGATCAGATTTGATCCCTGCGCCATCAGAT 4990
Db 3005 CATGCGAAACGATCCTCATCTCTGTCTTGTGATCAGAGTTTATCCCTGCGCCATCAGAT 2946
Qy 4991 CTTTGGCGGCAAGAAAGCATCCAGTTTACTTTCAGGGCTTCCCAACCTTACAGAGGG 5050
Db 2945 CTTTGGCGGCGAGAAAGCATCCAGTTTACTTTCAGGGCTTCCCAACCTTACAGAGGG 2886
Qy 5051 CGCCCCAGCTGGCAATTCGGGTTTGGCTTGTCTTGTGATCAATAAACCGCCAGTCTAGTATCG 5110
Db 2885 CGCCCCAGCTGGCAATTCGGGTTTGGCTTGTCTTGTGATCAATAAACCGCCAGTCTAGTATCG 2826
Qy 5111 CCATGTAAGCCCATCGAGCTACCTGCTTCTTTCGCTTGGCTTGGCTTTCCTTGTGCA 5170
Db 2825 CCATGTAAGCCCATCGAGCTACCTGCTTCTTTCGCTTGGCTTTCCTTGTGCA 2766
Qy 5171 GATAGCCAGTAGCTGACATTCATCCGGGGTTCAGCACCGTTTTCGCGACTGGCTTCTA 5230
Db 2765 GATAGCCAGTAGCTGACATTCATCCGGGGTTCAGCACCGTTTTCGCGACTGGCTTCTA 2706
Qy 5231 CGTG 5234
Db 2705 CGTG 2702

RESULT 6
US-10-021-403A-10
; Sequence 10, Application US/10021403A
; Publication No. US20030074679A1
; GENERAL INFORMATION:
; APPLICANT: Advisys
; TITLE OF INVENTION: Administration of Nucleic Acid Sequence to Female Animal to Enh
; FILE OF INVENTION: Growth in Offspring
; FILE REFERENCE: HO-P02021US1/100021476/OTA 00-91
; CURRENT APPLICATION NUMBER: US/10/021,403A
; PRIOR FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/255,021
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 2192
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: This is a plasmid vector comprising a pVC0289 backbone
US-10-021-403A-10

Query Match 30.4%; Score 1603.8; DB 15; Length 2192;
Best Local Similarity 82.8%; Pred. No. 0;
Matches 2022; Conservative 0; Mismatches 27; Indels 392; Gaps 1;

Qy 2706 CTAGAGCTAATCATGTGTATGCTGTTCTCTGTGTGAAATTTGTTATCCGCTCACAATTC 2765
Db 70 CTTTGGCGTAAATCATGTGTATGCTGTTCTCTGTGTGAAATTTGTTATCCGCTCACAATTC 129
Qy 2766 ACACAACATACGAGCCGGAAGCAATAAGTCTAAAGCTGGGTGCTTATGATGAGCTA 2825
Db 130 ACACAACATACGAGCCGGAAGCAATAAGTCTAAAGCTGGGTGCTTATGATGAGCTA 189
Qy 2826 ACTCAGATTAATTCGTTGGCTCACTGCCGCTTTCAGTCCGGAACCTGTCGTGCCA 2885
Db 190 ACTCAGATTAATTCGTTGGCTCACTGCCGCTTTCAGTCCGGAACCTGTCGTGCCA 249
Qy 2886 GCTGCATTAATTAATCGGCAACCGCGGAGAGCGGTTTCGCTATTGCGGCGCTTTC 2945
Db 250 GCTGCATTAATTAATCGGCAACCGCGGAGAGCGGTTTCGCTATTGCGGCGCTTTC 309
Qy 2946 CGTTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 3005
Db 310 CGTTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 369
```

3006	TC	ACTCAAAGCGCGTAA	TACGGTTAT	TCCAGAAATCAGGGGATAGCGACGAAAGAACAT	3065	
370	TC	ACTCAAAGCGCGTAA	TACGGTTAT	TCCAGAAATCAGGGGATAGCGACGAAAGAACAT	429	
3066	GT	CAGCAAAAGGCGCAGCA	MAAGGCGCAGGACCGTAA	AAAGGCGCGCTTCTGCGCGTTT	3125	
430	GT	CAGCAAAAGGCGCAGCA	MAAGGCGCAGGACCGTAA	AAAGGCGCGCTTCTGCGCGTTT	489	
3126	CC	ATAGGCTCCGCCGCCCT	GACGAGCATCA	AAAAATCGACGCTCAAGTCAGAGGTGGCG	3185	
490	CC	ATAGGCTCCGCCGCCCT	GACGAGCATCA	AAAAATCGACGCTCAAGTCAGAGGTGGCG	549	
3186	AA	ACCGCAGAGGACTATA	AAAGATAC	CAGCGGTTCCCTTGGAAGTCCTCTCGTCGCTC	3245	
550	AA	ACCGCAGAGGACTATA	AAAGATAC	CAGCGGTTCCCTTGGAAGTCCTCTCGTCGCTC	609	
3246	TC	CTGTTCCGACCGCTG	CGCGTTACCGGATAC	CTGTCGCGCTTCTCCCTTCGGGAAGCGT	3305	
610	TC	CTGTTCCGACCGCTG	CGCGTTACCGGATAC	CTGTCGCGCTTCTCCCTTCGGGAAGCGT	669	
3306	GG	CGCTTCTCATAGCTCA	CGCTGTAGGTATCT	CAGTTCGGTCTAGGTCTCGCTCCAA	3365	
670	GG	CGCTTCTCATAGCTCA	CGCTGTAGGTATCT	CAGTTCGGTCTAGGTCTCGCTCCAA	729	
3366	GC	TGGGCTGTGTGCAG	AAACCCCGCTT	CAGCCGACCGCTGCGCTTATCCGGTAACTA	3425	
730	GC	TGGGCTGTGTGCAG	AAACCCCGCTT	CAGCCGACCGCTGCGCTTATCCGGTAACTA	789	
3426	TC	GTCCTGAGTCCAA	CCCGGTAA	GACGAGCTTATCGCCTGCGCAGCAGCAGCTGGTAA	3485	
790	TC	GTCCTGAGTCCAA	CCCGGTAA	GACGAGCTTATCGCCTGCGCAGCAGCAGCTGGTAA	849	
3486	CAG	GATTAGCAGACGAG	GTATGAGCGGTG	TACAGAGTTCTTGAAGTGGTGGCTAA	3545	
850	CAG	GATTAGCAGACGAG	GTATGAGCGGTG	TACAGAGTTCTTGAAGTGGTGGCTAA	909	
3546	CT	ACGGCTACACTAGA	AAACAGTATTTGGTATCT	CGCTCTGCTGAGCAGCTTACCTT	3605	
910	CT	ACGGCTACACTAGA	AAACAGTATTTGGTATCT	CGCTCTGCTGAGCAGCTTACCTT	969	
3606	CG	AAAAAGAGTTGGT	AGCTCTTGATCCGGCAAA	CAAAACCAACCGCTGGTACGCGTGGTTT	3665	
970	CG	AAAAAGAGTTGGT	AGCTCTTGATCCGGCAAA	CAAAACCAACCGCTGGTACGCGTGGTTT	1029	
3666	TTTT	TTTTGTTG	ACGACAGATTA	CGCGCAGAAAAAAGGATCTCA	GAAAGATCTCTTGAT	3725
1030	TTTT	TTTTGTTG	ACGACAGATTA	CGCGCAGAAAAAAGGATCTCA	GAAAGATCTCTTGAT	1089
3726	CT	TTTCTACGGGGTCT	CAGCTCAGTGG	AACGAAACTCACGTTAAGGGATTTTGGTCA	3785	
1090	CT	TTTCTACGGGGTCT	CA	-----	1107	
3786	GAG	ATTATGTCGAC	CAAAAGCGGCGCAT	TCGTGCTCCCACTCTCTGCAATTCGGGGGCGATG	3845	
1108	-----	-----	-----	-----	1107	
3846	GAT	CGCGGATAGCGGT	CTGCTGTTCTCTGGAT	TCGACGAGGATTTGCACTGCCGATGAA	3905	
1108	-----	-----	-----	-----	1107	
3906	CT	CGCGAGGTGTC	ACGCTCA	GGCAGACGCTGAA	CAAACTCGCGAGGGGATCGAGCC	3965
1108	-----	-----	-----	-----	1107	
3966	GG	GTGGCGAAGAACT	CCAGATGAGAT	CCCCGCTGGAGGATCAT	CCAGCCGCGCTC	4025
1108	-----	-----	-----	-----	1107	
4026	CC	GAAAAACGATT	CCGAAGCCCA	ACCTTTTCATAG	AAGCGCGGTGGAAATCGAAATCTCG	4085
1108	-----	-----	-----	-----	1107	

QY	4086	TGATGGCAGGTTGGGGCGTTCGCTTGGTTCGGTCAATTTGGAACCCACAGAGTCCCGCTCAGAAG	41445
DB	1108	-----CGTCTCAGAAG	11117
QY	4146	AACTCGTCAAGAAAGCGGATAGAAGCGATTCGCTGCGAATTCGGAGCGGGCGATACCGGTAA	4205
DB	1118	AACTCGTCAAGAAAGCGGATAGAAGCGATTCGCTGCGAATTCGGAGCGGGCGATACCGGTAA	1177
QY	4206	AGCACGAGGAAGCGGTCAGGCCATTCGCGCGCAAGCTCTTCAGCAATATCACGGGTAGCC	4265
DB	1178	AGCACGAGGAAGCGGTCAGGCCATTCGCGCGCAAGCTCTTCAGCAATATCACGGGTAGCC	1237
QY	4266	AACGCTATGCTCTGATAGCGGTTCGGCCACACCCAGCGCGGCACAGTCGATGAATCCAGAA	4325
DB	1238	AACGCTATGCTCTGATAGCGGTTCGGCCACACCCAGCGCGGCACAGTCGATGAATCCAGAA	1297
QY	4326	AAGCGGCGATTTCCACCAATGATATTCGCGCAAGCAGGCATTCGCATTCGGTTCAGACGAGA	4385
DB	1298	AAGCGGCGATTTTCACCAATGATATTCGCGCAAGCAGGCATTCGCATTCGGTTCAGACGAGA	1357
QY	4386	TCCTTCGCGCGTCGGGCATTCGCGCTTGAGCCTTCGGCGAACAGTTTCGGCTGCGCGAGCCCC	4445
DB	1358	TCCTTCGCGCGTCGGGCATTCGCGCTTGAGCCTTCGGCGAACAGTTTCGGCTGCGCGAGCCCC	1417
QY	4446	TGATGCTCTTCGTCAGATCATCTGATTCAGACAGCAGGCATTCGCATTCGGTTCAGATCGTGCT	4505
DB	1418	TGATGCTCTTCGTCAGATCATCTGATTCAGACAGCAGGCATTCGCATTCGGTTCAGATCGTGCT	1477
QY	4506	CGCTTCGATGCGATGTTTCGCTTCGTTGCTGCTGCAATTCGGCAGGTAGCCGATCAAGCGTATGC	4565
DB	1478	CGCTTCGATGCGATGTTTCGCTTCGTTGCTGCTGCAATTCGGCAGGTAGCCGATCAAGCGTATGC	1537
QY	4566	AGCGCGCGCATTCGATCAGCCATGATCGATTCCTTCGCGAGGACCGGCTTCATTCGAGTACGTCG	4625
DB	1538	AGCGCGCGCATTCGATCAGCCATGATCGATTCCTTCGCGAGGACCGGCTTCATTCGAGTACGTCG	1597
QY	4626	AGGAGATCCTGCCCGGCACCTTCGCCCAATAGCAGCCAGTCCCTTCGCGTTCAGTGACA	4685
DB	1598	AGGAGATCCTGCCCGGCACCTTCGCCCAATAGCAGCCAGTCCCTTCGCGTTCAGTGACA	1657
QY	4686	ACGTCGAGCACAGCTGCGCAGGAAGACCCGCTGCTGCGCAGCAGATAGCCGCGCTGCC	4745
DB	1658	ACGTCGAGCACAGCTGCGCAGGAAGACCCGCTGCTGCGCAGCAGATAGCCGCGCTGCC	1717
QY	4746	TCGTCCTCGAGTTCATTAAAGGGCACCGGACAGGTCGCTCTTGACAAAAGAACCGGGGCGC	4805
DB	1718	TCGTCCTCGAGTTCATTAAAGGGCACCGGACAGGTCGCTCTTGACAAAAGAACCGGGGCGC	1777
QY	4806	CCCTTCGCTGACACGCGGACACGCGCGCATTCAGACAGCGGATGCTGTTGTCGCCAG	4865
DB	1778	CCCTTCGCTGACACGCGGACACGCGCGCATTCAGACAGCGGATGCTGTTGTCGCCAG	1837
QY	4866	TCATAGCGGAATAGCCTCTCCACCCAGCGCGCGGAGAACTCGGCTGGAAATCCGATCTGT	4925
DB	1838	TCATAGCGGAATAGCCTCTCCACCCAGCGCGCGGAGAACTCGGCTGGAAATCCGATCTGT	1897
QY	4926	TCAATCATGCGAAACGATCCTCATCTCTGCTCTTTGATTCAGATCTTGATCCCTCGCGCAT	4985
DB	1898	TCAATCATGCGAAACGATCCTCATCTCTGCTCTTTGATTCAGATCTTGATCCCTCGCGCAT	1957
QY	4986	CAGATCTCTGGCGGAAGAAAGCATTCAGTTTACTTTGACGGGCTTCGCAACCTTACCA	5045
DB	1958	CAGATCTCTGGCGGAAGAAAGCATTCAGTTTACTTTGACGGGCTTCGCAACCTTACCA	2017
QY	5046	GAGGGCGGCCAGCTGGGCAATTCGGGTTTCGCTGCTCTGCTGCTCATATAAAACCGCCAGCTTAGC	5105
DB	2018	GAGGGCGGCCAGCTGGGCAATTCGGGTTTCGCTGCTCTGCTGCTCATATAAAACCGCCAGCTTAGC	2077
QY	5106	TATCGCCATGTAAGCCCACTCGAAGCTTACCTGCTTCTCTT	5146
DB	2078	AACTCTTGGGAAGGCGCATTCGGTCGGGCGCTCTTCGCTATT	2118

RESULT 7

RECORD
 US-10-395-709-11
 ; Sequence 11, Application US/10395709
 ; Publication No. US20040014645A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Advivys
 ; TITLE OF INVENTION: INCREASED DELIVERY OF A NUCLEIC ACID CONSTRUCT IN VIVO BY THE POL
 ; TITLE OF INVENTION: GLUTAMATE ("PLIG") SYSTEM
 ; FILE REFERENCE: 108328.00115 - AVSI-0021P1
 ; CURRENT APPLICATION NUMBER: US/10/395,709
 ; CURRENT FILING DATE: 2003-03-24
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 11
 ; LENGTH: 3534
 ; TYPE: DNA
 ; ORGANISM: artificial sequence
 ; FEATURE:
 ; OTHER INFORMATION: Sequence for the operatively linked components of the HV-GHRH pla
 ; OTHER INFORMATION: smid.
 US-10-395-709-11

Query Match 30.4%; Score 1603.8; DB 12; Length 3534;
Best Local Similarity 82.8%; Pred. No. 0;
Matches 2022; Conservative 0; Mismatches 27; Indels 392; Gaps 1;

Qy	2706	CTAGACGTAATCATGGTTCATAGCTGTTTCTGTGTGAATTTGTTATCCGCTCAATTC	2765
Db	1412	CTTGGCGTAATCATGGTTCATAGCTGTTTCTGTGTGAATTTGTTATCCGCTCAATTC	1471
Qy	2766	ACAAACATACGAGCCGGAAGCATAAAGTGTAAAGCCCTGGGGTGCCTAATAGAGTGAGCTA	2825
Db	1472	ACAAACATACGAGCCGGAAGCATAAAGTGTAAAGCCCTGGGGTGCCTAATAGAGTGAGCTA	1531
Qy	2826	ACTCACATTAATTGCGTTTGCGCTCAGTCGCCGCTTTCAGTTCGGGAAACCTGTCTGCGCA	2885
Db	1532	ACTCACATTAATTGCGTTTGCGCTCAGTCGCCGCTTTCAGTTCGGGAAACCTGTCTGCGCA	1591
Qy	2886	GCTGCATTAAATGAAATCGGCCAAACGCGCGGGAGAGCGGTTTGCGTATTGGGCGCTCTTC	2945
Db	1592	GCTGCATTAAATGAAATCGGCCAAACGCGCGGGAGAGCGGTTTGCGTATTGGGCGCTCTTC	1651
Qy	2946	CGCTTCCTCGTCACTGACTCGCTGCGCTCGGTTCGGTTCGGGTCGGGCGAGCGGTATCAGC	3005
Db	1652	CGCTTCCTCGTCACTGACTCGCTGCGCTCGGTTCGGTTCGGGTCGGGCGAGCGGTATCAGC	1711
Qy	3006	TCACCTCAAGCGCGGTAAATACGTTTATCCACAGAAATCAGGGGATAAACGACGGAAGAACAT	3065
Db	1712	TCACCTCAAGCGCGGTAAATACGTTTATCCACAGAAATCAGGGGATAAACGACGGAAGAACAT	1771
Qy	3066	GTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGCCGCGTTCGTGCGGTTTTT	3125
Db	1772	GTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGCCGCGTTCGTGCGGTTTTT	1831
Qy	3126	CCATAGGCTCGCGCCCTGACGAGCATCAAAAATTCGACGCTCAAGTCAGAGGTGGCG	3185
Db	1832	CCATAGGCTCGCGCCCTGACGAGCATCAAAAATTCGACGCTCAAGTCAGAGGTGGCG	1891
Qy	3186	AAACCCGACAGGACTATAAAGATACCAAGCGCTTTCCCGCTGGAAAGTCCCTCTGTCGCGCTC	3245
Db	1892	AAACCCGACAGGACTATAAAGATACCAAGCGCTTTCCCGCTGGAAAGTCCCTCTGTCGCGCTC	1951
Qy	3246	TCCTGTTCCGACCTGCGGCTTACGGATACCTGTGCGCTTTCTCCGTTTCGGGAAGCGT	3305
Db	1952	TCCTGTTCCGACCTGCGGCTTACGGATACCTGTGCGCTTTCTCCGTTTCGGGAAGCGT	2011
Qy	3306	GGCGCTTTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTTCGTCGCTCAA	3365
Db	2012	GGCGCTTTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTTCGTCGCTCAA	2071
Qy	3366	GCTGGGCTGTGTGACGAACCCCGGTTTCAGCCCGAACGCGTCGCGCTTATCCGGTAACATA	3425
Db	2072	GCTGGGCTGTGTGACGAACCCCGGTTTCAGCCCGAACGCGTCGCGCTTATCCGGTAACATA	2131

QY 3846 GATCGCGGATAGCCGCTGCTGCTTTCTGTGATGCGGACGGATTGCACTGCGCGTAGAA 3905
Db 2450 ----- 2449
QY 3906 CTCCGCGAGTGGTCCAGCCTCAGGACGAGCTGACCACTCGCGAGGGGATCGAGCCC 3965
Db 2450 ----- 2449
QY 3966 GGGGTGGGCGAAGAACTCCAGCATGAGATCCCGCGCTGGAGGATCATCCAGCCGCGTC 4025
Db 2450 ----- 2449
QY 4026 CCGGAAAACGATTCCGAAGCCCACTTTTCATAGAAGCGCGGTGGGAATCGAATCTCG 4085
Db 2450 ----- 2449
QY 4086 TGATGCGAGTGGGCGTGGCTTGGTGGTCTGATTTTCGAACCCAGAGTCCCGCTCAGAAG 4145
Db 2450 -----CGCTCAGAAG 2459
QY 4146 AACTCGTCAAGAGGCGATAGAGGCGATGCGCTGCGAATCGGAGCGGCGATACCGTAA 4205
Db 2460 AACTCGTCAAGAGGCGATAGAGGCGATGCGCTGCGAATCGGAGCGGCGATACCGTAA 2519
QY 4206 AGCAGGAGGAGCGGTCAGCCATTCGCGCGCAAGCTCTTCAGCAATATCACGGGTAGCC 4265
Db 2520 AGCAGGAGGAGCGGTCAGCCATTCGCGCGCAAGCTCTTCAGCAATATCACGGGTAGCC 2579
QY 4266 AACGCTATGCTCTGATAGCGGTCCGCCACACCCAGCCGCGCATGCGATGATGATCCAGAA 4325
Db 2580 AACGCTATGCTCTGATAGCGGTCCGCCACACCCAGCCGCGCATGCGATGATGATCCAGAA 2639
QY 4326 AAGCGGCCATTTCCACCATGATATTCGGAAGCAGGCGATCGCATGGTCAAGCAGAGA 4385
Db 2640 AAGCGGCCATTTCCACCATGATATTCGGAAGCAGGCGATCGCATGGTCAAGCAGAGA 2699
QY 4386 TCCTCGCGCTCGGCGCATGCGGCTTGGCGAAGCAGTTTGGCTGGCGCGGCGAGCCCC 4445
Db 2700 TCCTCGCGCTCGGCGCATGCGGCTTGGCGAAGCAGTTTGGCTGGCGCGGAGCCCC 2759
QY 4446 TGATGCTCTGCTCGAGATCATCTGATCGACAGACCGGCTTCATTCGAGTACGTGCT 4505
Db 2760 TGATGCTCTGCTCGAGATCATCTGATCGACAGACCGGCTTCATTCGAGTACGTGCT 2819
QY 4506 CGCTCGATGCGATGTTTGGCTTGGTTCGATGCGGAGGAGTACGCGGATCAAGCGTATGC 4565
Db 2820 CGCTCGATGCGATGTTTGGTTCGATGCGGAGGAGTACGCGGATCAAGCGTATGC 2879
QY 4566 AGCCGCGCATTCGATCAGCCATGATGGATATCTTTCTCGGCGAGGAGCAAGGTGATGAC 4625
Db 2880 AGCCGCGCATTCGATCAGCCATGATGGATATCTTTCTCGGCGAGGAGCAAGGTGATGAC 2939
QY 4626 AGGAGATCTGCGCGGCACTTCGCCCAATAGCAGCGCTGCTTCCCGCTTCAGTGACA 4685
Db 2940 AGGAGATCTGCGCGGCACTTCGCCCAATAGCAGCGCTGCTTCCCGCTTCAGTGACA 2999
QY 4686 ACCTCGAGCACAGCTGCGCAAGGAAGCCCGTCTGTCGCGCAGGACGATAGCCGCGTGC 4745
Db 3000 ACCTCGAGCACAGCTGCGCAAGGAAGCCCGTCTGTCGCGCAGGACGATAGCCGCGTGC 3059
QY 4746 TCGTCTGCGATGCTATTCAGGCGACCGGACAGGTGCTTTCGACAAAAGAACCGGGGCG 4805
Db 3060 TCGTCTGCGATGCTATTCAGGCGACCGGACAGGTGCTTTCGACAAAAGAACCGGGGCG 3119
QY 4806 CCCTGCGCTGACAGCGGAAACCGGCGGATCAGAGCAGCCGATTTGCTGTTGCGCCAG 4865
Db 3120 CCCTGCGCTGACAGCGGAAACCGGCGGATCAGAGCAGCCGATTTGCTGTTGCGCCAG 3179
QY 4866 TCATAGCCGAATAGCTTTCACCCCAAGCGGCGGAGAACCTTGGCTGCGAATCCATCTTGT 4925
Db 3180 TCATAGCCGAATAGCTTTCACCCCAAGCGGCGGAGAACCTTGGCTGCGAATCCATCTTGT 3239

QY 4926 TCAATCATGCGAAACGATCCTCATCTCTGTCTCTTGATCAGATCTTGATCCCTCGCGCAT 4985
Db 3240 TCAATCATGCGAAACGATCCTCATCTCTGTCTCTTGATCAGATCTTGATCCCTCGCGCAT 3299
QY 4986 CAGATCCTTCGCGCAAGAACGATCCAGTTTACTTTTTCAGGGCTTCCCAACCTTACCA 5045
Db 3300 CAGATCCTTCGCGCAAGAACGATCCAGTTTACTTTTTCAGGGCTTCCCAACCTTACCA 3359
QY 5046 GAGGCGCCCGCAGCTGCGAATTCGGTTCGCTTGTGTCATATAAACCGCCAGTCTAGC 5105
Db 3360 GAGGCGCCCGCAGCTGCGAATTCGGTTCGCTTGTGTCATATAAACCGCCAGTCTAGC 3419
QY 5106 TATCGCCATGTAAGCCCACTGCAAGTACCTGCTTTCTCTT 5146
Db 3420 AACTGTTGGAGGCGATCGGTGCGGCGCTCTTCGCTATT 3460

RESULT 9
US-10-395-709-13
; Sequence 13, Application US/10395709
; Publication No. US20040014645A1
; GENERAL INFORMATION:
; APPLICANT: ADVISYS
; TITLE OF INVENTION: INCREASED DELIVERY OF A NUCLEIC ACID CONSTRUCT IN VIVO BY THE F
; FILE REFERENCE: 108328.00115 - AVSI-0021P1
; CURRENT APPLICATION NUMBER: US/10395,709
; CURRENT FILING DATE: 2003-03-24
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 3534
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Sequence for the operatively linked components of the TV-GHRH F
; OTHER INFORMATION: smid.
US-10-395-709-13

Query Match 30.4%; Score 1603.8; DB 12; Length 3534;
Best Local Similarity 82.8%; Pred. No. 0;
Matches 2022; Conservative 0; Mismatches 27; Indels 392; Gaps 1;
QY 2706 CTAGACGTAATCATGCTAGCTGTTTCTGTTGTAATTTGTTATCCGCTCAATTC 2765
Db 1412 CTGCGCTAATCATGCTAGCTGTTTCTGTTGTAATTTGTTATCCGCTCAATTC 1471
QY 2766 ACACAAATAGAGCCGGAAGCATTAAGTCTAAGCTGCGGTGCTTAATGAGTGAGCTA 2825
Db 1472 ACACAAATAGAGCCGGAAGCATTAAGTCTAAGCTGCGGTGCTTAATGAGTGAGCTA 1531
QY 2826 ACTCACATTAATTCGCTTGGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCA 2885
Db 1532 ACTCACATTAATTCGCTTGGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCA 1591
QY 2886 GCTGCATTAATGATCGGCCAAGCGCGGGAGAGCGGTTTTCGTTATTCGGGCGCTTTC 2945
Db 1592 GCTGCATTAATGATCGGCCAAGCGCGGGAGAGCGGTTTTCGTTATTCGGGCGCTTTC 1651
QY 2946 CGCTTCTCGCTCACTCACTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTATCAGC 3005
Db 1652 CGCTTCTCGCTCACTCACTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTATCAGC 1711
QY 3006 TCATCAAAAGCGGTAATACGGTTATCCACAGAAATCAGGGGATAACCGAGAAAGAACAT 3065
Db 1712 TCATCAAAAGCGGTAATACGGTTATCCACAGAAATCAGGGGATAACCGAGAAAGAACAT 1771
QY 3066 GTGAGCAAAAGCGGAGCAAGGAGCGGAGGAGCGGTTGTTGCGGTTTTC 3125
Db 1772 GTGAGCAAAAGCGGAGCAAGGAGCGGAGGAGCGGTTGTTGCGGTTTTC 1831
QY 3126 CCATAGGCTCGGCCCTTCGACGAGCATCAAAAATTCAGCTCAAGTCAGAGTGGCG 3185

Db 1832 CCAATAGGCTCGCCCCCTGACGAGCATCAAAAAATCGAGCTCAAGTCAGAGGTGGCG 1891
Qy 3186 AAACCCGACAGGACTATAAGATACACAGGGGTTTCCCCCTGGAGAGTCTCCCTCGTGGCTC 3245
Db 1892 AAACCCGACAGGACTATAAGATACACAGGGGTTTCCCCCTGGAGAGTCTCCCTCGTGGCTC 1951
Qy 3246 TCCTGTTCCGACCTCGCGCTTACCGATACCTGTCGCGCTTCTCCCTTCGGGAGCGT 3305
Db 1952 TCCTGTTCCGACCTCGCGCTTACCGATACCTGTCGCGCTTCTCCCTTCGGGAGCGT 2011
Qy 3306 GCGCTTCTCATAGCTCACGCTGAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAA 3365
Db 2012 GCGCTTCTCATAGCTCACGCTGAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAA 2071
Qy 3366 GCTGGGCTGTGCAAGAACCCCGCTTACGCGCAGCGCTTACCGCTTATCCGGTAAC 3425
Db 2072 GCTGGGCTGTGCAAGAACCCCGCTTACGCGCAGCGCTTATCCGGTAAC 2131
Qy 3426 TCCTGTTGAGTCCAAACCGGTAAACACAGCTTATCGCCACTGCGACAGCCACTGTAA 3485
Db 2132 TCGTCTTGAGTCCAAACCGGTAAACACAGCTTATCGCCACTGCGACAGCCACTGTAA 2191
Qy 3486 CAGGATTAGCAGCGAGGTATGAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTAA 3545
Db 2192 CAGGATTAGCAGCGAGGTATGAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTAA 2251
Qy 3546 CTAGCGCTACACTAGAGACAGTATTTGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3605
Db 2252 CTAGCGCTACACTAGAGACAGTATTTGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2311
Qy 3606 CGGAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACACACCGCTGGTAGCGGTGTT 3665
Db 2312 CGGAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACACACCGCTGGTAGCGGTGTT 2371
Qy 3666 TTTTGTGTTGAAAGCAGATAGCGGAGAAAAGAGTCTCAAGAGATCCTTTGAT 3725
Db 2372 TTTTGTGTTGAAAGCAGATAGCGGAGAAAAGAGTCTCAAGAGATCCTTTGAT 2431
Qy 3726 CTTTCTACGGGGTCTGACGCTCAGTGGAGCAAAACCTACGTTAAGGATTTTGGTCAT 3785
Db 2432 CTTTCTACGGGGTCTGACGCTCAGTGGAGCAAAACCTACGTTAAGGATTTTGGTCAT 3845
Qy 3786 GAGATTATGTCGACCAAGCGGCGCATGCTGCTCCCACTCTCTGCAAGTTTCGGGGGCGATG 3845
Db 2450 ----- 2449
Qy 3846 GATCGGAGATAGCGCTGCTGCTGTTTCTGATGCGGAGGATTTGCACTGCGGCTAGAA 3905
Db 2450 ----- 2449
Qy 3906 CTCCGCGAGTGTCTCAGCTTCAGGACAGCTGAACCACTCGCGAGGGGATCGAGCC 3965
Db 2450 ----- 2449
Qy 3966 GGGGTGGGGAAGAACTCCAGATGAGATCCCGCGCTGGAGGATCATCAGCGGGCTC 4025
Db 2450 ----- 2449
Qy 4026 CCGGAAAACGATTCGGAAGCCCAACTTTTCATAGAGCGCGGTGGAATCGAAATCTCG 4085
Db 2450 ----- 2449
Qy 4086 TGATGGCAGGTTGGCGCTCGCTTGGTGGTCAATTCGAACCCAGAGTCCCGCTCAGAAG 4145
Db 2450 -----CGCTCAGAAG 2459
Qy 4146 AACTCGTCAAGAAGCGATAGAGCGATGGCTGCGAATCGGAGCGGCGATACCGTAA 4205
Db 2460 AACTCGTCAAGAAGCGATAGAGCGATGGCTGCGAATCGGAGCGGCGATACCGTAA 2519
Qy 4206 AGCAGAGAGCGGTGAGCCATTCGCGGCAAGCTCTTCAGCAATATCAGGGGTAGCC 4265
Db 2520 AGCAGAGAGCGGTGAGCCATTCGCGGCAAGCTCTTCAGCAATATCAGGGGTAGCC 2579

Qy 4266 AACGCTATGCTCTGATAGCGTCCGCCACACCCAGCGGCCACAGTCGATGAATCCAGAA 4325
Db 2580 AACGCTATGCTCTGATAGCGTCCGCCACACCCAGCGGCCACAGTCGATGAATCCAGAA 2639
Qy 4326 AAGCGCCATTTTCCACATGATATTGGCAAGCAGGATCGCCATCGGTTCACGACGAGA 4385
Db 2640 AAGCGCCATTTTCCACATGATATTGGCAAGCAGGATCGCCATCGGTTCACGACGAGA 2699
Qy 4386 TCCTCGCGCTCGGCGATCGCGCTTGAAGCTTGGCGAACAGTTCGGTGGCGGAGGCC 4445
Db 2700 TCCTCGCGCTCGGCGATCGCGCTTGAAGCTTGGCGAACAGTTCGGTGGCGGAGGCC 2759
Qy 4446 TGATGCTCTTCTGTCAGATCATCTGATCGACAGACCGGCTTCATCCGAGTACGTCT 4505
Db 2760 TGATGCTCTTCTGTCAGATCATCTGATCGACAGACCGGCTTCATCCGAGTACGTCT 2819
Qy 4506 CGCTCGATGCGATGTTTCGCTTGGTGGTTCGAATGGCGAGGTAGCGGATCAAGCGTATGC 4565
Db 2820 CGCTCGATGCGATGTTTCGCTTGGTGGTTCGAATGGCGAGGTAGCGGATCAAGCGTATGC 2879
Qy 4566 AGCGCCCGATTCGATCGACCATGATGATCTTCTCGGAGGAGGAGTGGATGAC 4625
Db 2880 AGCGCCCGATTCGATCGACCATGATGATCTTCTCGGAGGAGGAGTGGATGAC 2939
Qy 4626 AGGAGATCCTGCGCGGCACTTCGCCAATAGCAGCCAGTCCCTTCCCGCTTCAGTGACA 4685
Db 2940 AGGAGATCCTGCGCGGCACTTCGCCAATAGCAGCCAGTCCCTTCCCGCTTCAGTGACA 2999
Qy 4686 AGCTCGACACAGCTGCGCAAGAAAGCGCCGCTGCTGGCAGCAGATAGCCGCGCTGCC 4745
Db 3000 AGCTCGACACAGCTGCGCAAGAAAGCGCCGCTGCTGGCAGCAGATAGCCGCGCTGCC 3059
Qy 4746 TCGTCTCTCAGTTCATTCAGGCGCACCGGACAGTTCGTTGACAAAAGAACCGGGCGC 4805
Db 3060 TCGTCTCTCAGTTCATTCAGGCGCACCGGACAGTTCGTTGACAAAAGAACCGGGCGC 3119
Qy 4806 CCCTCGCTGACAGCGGAAACACCGCGGATCAGAGCAGCGGATGCTGTTGTGCCAG 4865
Db 3120 CCCTCGCTGACAGCGGAAACACCGCGGATCAGAGCAGCGGATGCTGTTGTGCCAG 3179
Qy 4866 TCATAGCGGATAGCTTCCACCGAGCGCGGAGACCTGCTGCTCAATCCATCTGT 4925
Db 3180 TCATAGCGGATAGCTTCCACCGAGCGCGGAGACCTGCTGCTCAATCCATCTGT 3239
Qy 4926 TCAATATCGGAAAGCATCTCATCTCTCTTGTATCAGATCTTGATCCCTCGCCAT 4985
Db 3240 TCAATATCGGAAAGCATCTCATCTCTCTTGTATCAGATCTTGATCCCTCGCCAT 3299
Qy 4986 CAGATCTTGGCGGCAAGAAAGCATCCAGTTTACTTTGAGGGCTTCCGACCTTACCA 5045
Db 3300 CAGATCTTGGCGGCAAGAAAGCATCCAGTTTACTTTGAGGGCTTCCGACCTTACCA 3359
Qy 5046 GAGGCGCGCCAGCTGGCAATTCGGTTCGCTTCTGCTCAATAAAACCGCCAGTCTAGC 5105
Db 3360 GAGGCGCGCCAGCTGGCAATTCGGTTCGCTTCTGCTTCCATATAACCGCCAGTCTAGC 3419
Qy 5106 TATGCCCATGTAAGCCCACTGCAAGTACCTGCTTCTCTT 5146
Db 3420 AACTGTTGGAAGGCGGATCGGTGCGGCTCTTCGCTATT 3460

RESULT 10

US-10-395-709-14
; Sequence 14, Application US/10395709
; Publication No. US20040014645A1
; GENERAL INFORMATION:
; APPLICANT: Advisys
; TITLE OF INVENTION: INCREASED DELIVERY OF A NUCLEIC ACID CONSTRUCT IN VIVO BY THE
; TITLE OF INVENTION: GLUTAMATE ("PLG") SYSTEM
; FILE REFERENCE: 108328.00115 - AVSI-0021P1
; CURRENT APPLICATION NUMBER: US/10/395,709
; CURRENT FILING DATE: 2003-03-24

; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 14
; LENGTH: 3534
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Sequence for the operatively linked components of the 15/27/28 GH
; OTHER INFORMATION: RH plasmid.
US-10-395-709-14

Query Match 30.4%; Score 1603.8; DB 12; Length 3534;
Best Local Similarity 82.8%; Pred. No. 0;
Matches 2022; Conservative 0; Mismatches 27; Indels 392; Gaps 1;
QY 2706 CTAGAGCTAATCATGTCTAGCTGTTCTCTGTGTAATTTATTCGCTCACAATTC 2765
Db 1412 CTTGGCGTAATCATGTCTAGCTGTTCTCTGTGTAATTTATTCGCTCACAATTC 1471
QY 2766 ACACACATACGAGCGGAGCATTAAGTGAAGCTGGGCTGCTTAATGATGAGCTA 2825
Db 1472 ACACACATACGAGCGGAGCATTAAGTGAAGCTGGGCTGCTTAATGATGAGCTA 1531
QY 2826 ACTCATTAAATTCGCTGCTCACTGCTCCGCTTTCCAGTCGGAAACCTGCTGCTA 2885
Db 1532 ACTCATTAAATTCGCTGCTCACTGCTCCGCTTTCCAGTCGGAAACCTGCTGCTA 1591
QY 2886 GCTGCATTAAATGATCGGCAACGCGGGGAGAGCGGTTTGGCTATTTGGCGCTCTTC 2945
Db 1592 GCTGCATTAAATGATCGGCAACGCGGGGAGAGCGGTTTGGCTATTTGGCGCTCTTC 1651
QY 2946 CGCTTCTCGCTCACTGACTCGCTGCGCTCGGTGCTGCGCTGCGCGAGCGGTATCAGC 3005
Db 1652 CGCTTCTCGCTCACTGACTCGCTGCGCTCGGTGCTGCGCTGCGCGAGCGGTATCAGC 1711
QY 3006 TCACTCAAGCGGTAATACGTTATCCACAGATCAGGGGATACGCGAGGAAGAAAT 3065
Db 1712 TCACTCAAGCGGTAATACGTTATCCACAGATCAGGGGATACGCGAGGAAGAAAT 1771
QY 3066 GTGAGCAAAAGGCCACGAAAGCCAGGACCGTAAAGGCGCGTTCGTCGGCTTTT 3125
Db 1772 GTGAGCAAAAGGCCACGAAAGCCAGGACCGTAAAGGCGCGTTCGTCGGCTTTT 1831
QY 3126 CCATAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGTGGCG 3185
Db 1832 CCATAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGTGGCG 1891
QY 3186 ABAACCGACAGGACTATAAGATACGAGCGTTTCCCTCTGAAAGCTCCCTGCGCTC 3245
Db 1892 ABAACCGACAGGACTATAAGATACGAGCGTTTCCCTCTGAAAGCTCCCTGCGCTC 1951
QY 3246 TCCTGTTCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGT 3305
Db 1952 TCCTGTTCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGT 2011
QY 3306 GGGCTTCTCATAGTCAAGTGTAGGTATCTCAGTTCGGTGTAGTGGTTCGCTCAA 3365
Db 2012 GGGCTTCTCATAGTCAAGTGTAGGTATCTCAGTTCGGTGTAGTGGTTCGCTCAA 2071
QY 3366 GCTGGGCTGTGTGACGAAACCCCGCTTCAGCCCGACCGCTGCGCTTATCCGGTAACTA 3425
Db 2072 GCTGGGCTGTGTGACGAAACCCCGCTTCAGCCCGACCGCTGCGCTTATCCGGTAACTA 2131
QY 3426 TCGTCTTGAAGTCAACCCCGTAAAGACGACTATTCGCACTGGCAGAGCCACTGGTAA 3485
Db 2132 TCGTCTTGAAGTCAACCCCGTAAAGACGACTATTCGCACTGGCAGAGCCACTGGTAA 2191
QY 3486 CAGGATACGAGCGAGGTATGTAGCGGTCTACAGATTTCTTGAAGTGTGCGCTAA 3545
Db 2192 CAGGATACGAGCGAGGTATGTAGCGGTCTACAGATTTCTTGAAGTGTGCGCTAA 2251
QY 3546 CTACGGCTACCTAGAGAAACAGTATTTGGTATCTCGCTCTGCTGAAAGCCAGTACCTT 3605
Db 3546 CTACGGCTACCTAGAGAAACAGTATTTGGTATCTCGCTCTGCTGAAAGCCAGTACCTT 3605

Db 2252 CTACGGCTACCTAGAGAAACAGTATTTGGTATCTCGCTCTGCTGAAAGCCAGTTACCTT 2311
QY 3606 CGGAAAAAGAGTTGGTAGCTCTTTGATCCGGCAAAACAAACACCGCTGCTAGCGGTGTTT 3665
Db 2312 CGGAAAAAGAGTTGGTAGCTCTTTGATCCGGCAAAACAAACACCGCTGCTAGCGGTGTTT 2371
QY 3666 TTTTGTGTTGCAAGCAGAGATTAAGCGGCGAGAAAAAGGATCTCAAGAAAGATCTTTGAT 3725
Db 2372 TTTTGTGTTGCAAGCAGAGATTAAGCGGCGAGAAAAAGGATCTCAAGAAAGATCTTTGAT 2431
QY 3726 CTTTCTTACCGGGTCTGACGCTCAGTGTGAACGAAACTCAGCTTAAGGGATTTTGGTCAT 3785
Db 2432 CTTTCTTACCGGGTCTGACGCTCAGTGTGAACGAAACTCAGCTTAAGGGATTTTGGTCAT 2449
QY 3786 GAGATTATCTGACCAAAAGCGGCATCTGTCCTCCCACTCCTCTGAGTTTCGGGGGCAATG 3845
Db 2450 ----- 2449
QY 3846 GATCGCGGATAGCCGCTGCTGTTCTCTGGATGCCGACGGATTTGCACCTGCCGTAGAA 3905
Db 2450 ----- 2449
QY 3906 CTCGCGAGGTCTGTCAGCCTCAGGAGCAGCTGAACCACTCCGAGGGGATCGAGCCC 3965
Db 2450 ----- 2449
QY 3966 GGGTGGGCGAAGAACTCCAGCATGAGATCCCGCGCTGGAGGATCATCCAGCGCGGCTC 4025
Db 2450 ----- 2449
QY 4026 CCGAAAAAGATTCCGAAAGCCCAACTTTCTAGAAAGCGCGGTGGAATCGAAATCTCG 4085
Db 2450 ----- 2449
QY 4086 TGATGGCAGGTGGGCGTCTGCTTGGTGGTCAATTTGAAACCCAGAGTCCCGCTCAGAAG 4145
Db 2450 ----- 2459
QY 4146 AACTCGTCAAGAAAGCGGATAGAGGCGATGCGCTGCGAATCGGGAGCGCGATACCGTAA 4205
Db 2460 AACTCGTCAAGAAAGCGGATAGAGGCGATGCGCTGCGAATCGGGAGCGCGATACCGTAA 2519
QY 4206 AGCAGAGGAGCGGTGAGCCCACTTCGCGCCAAAGCTCTTCAGCAATATCAAGGCTAGCC 4265
Db 2520 AGCAGAGGAGCGGTGAGCCCACTTCGCGCCAAAGCTCTTCAGCAATATCAAGGCTAGCC 2579
QY 4266 AAGCTATGCTCTGATAGCGGTGCGGCACACCCAGCGGCGCAGTCGATGATCCAGAA 4325
Db 2580 AAGCTATGCTCTGATAGCGGTGCGGCACACCCAGCGGCGCAGTCGATGATCCAGAA 2639
QY 4326 AAGCGGCAATTTTCCACCATGATATTTCGGCAAGCAGGCATCGCCATGGGTACGACGAGA 4385
Db 2640 AAGCGGCAATTTTCCACCATGATATTTCGGCAAGCAGGCATCGCCATGGGTACGACGAGA 2699
QY 4386 TCCTTCGCGCTGCGGCAATGCGCGCTTGAAGCTGGGCAAGTTCGGCTGGCGGAGCGCC 4445
Db 2700 TCCTTCGCGCTGCGGCAATGCGCGCTTGAAGCTGGGCAAGTTCGGCTGGCGGAGCGCC 2759
QY 4446 TGATGCTCTTCGTCAGATCATCTGATCGAAGACCGGCTTCATCCGAGTACGTCGT 4505
Db 2760 TGATGCTCTTCGTCAGATCATCTGATCGAAGACCGGCTTCATCCGAGTACGTCGT 2819
QY 4506 CGCTCGATGCGATTTTTCGCTTGGTGGTGAATGGGAGGTAGCGATCCGATCAAGCGTATGC 4565
Db 2820 CGCTCGATGCGATTTTTCGCTTGGTGGTGAATGGGAGGTAGCGATCCGATCAAGCGTATGC 2879
QY 4566 AGCGCGCATTCGATCAGCATGATGAGTACTTTCTCGGAGGAGCAAGGTAGATGAC 4625
Db 2880 AGCGCGCATTCGATCAGCATGATGAGTACTTTCTCGGAGGAGCAAGGTAGATGAC 2939
QY 4626 AGGAGATCTTCGCGCGCGCATTCGCGCAATAGCAGCAGTCCCTTCCCGCTTCAGTGACA 4685
Db 2940 AGGAGATCTTCGCGCGCGCATTCGCGCAATAGCAGCAGTCCCTTCCCGCTTCAGTGACA 2999

QY 4026 CCGGAAACGATTCGGAAGCCCAACCTTTTCATAGAAGCGCGGTGGAATCGAAATCTCG 4085
Db 2450 ----- 2449
QY 4086 TGATGGCAGGTGGCGCTCGCTTGGTTCGGTCAATTCGAACCCAGAGTCCCGCTCAGAAG 4145
Db 2450 -----CGCTCAGAAG 2459
QY 4146 AACTCGTCAAGAAGCGGATAGAAGGCGATCGCTGCGAATCGGAGCGGCGATACCGTAA 4205
Db 2460 AACTCGTCAAGAAGCGGATAGAAGGCGATCGCTGCGAATCGGAGCGGCGATACCGTAA 2519
QY 4206 AGCAGAGGAGCGGTACGCCCATTCGCCGCGAAGCTCTTCAGCAATATACGGGTAGCC 4265
Db 2520 AGCAGAGGAGCGGTACGCCCATTCGCCGCGAAGCTCTTCAGCAATATACGGGTAGCC 2579
QY 4266 AAGCGTATGTCCTGATAGCGGTCCGCCACACCCAGCGCGCCACAGTCGATGAATCCAGAA 4325
Db 2580 AAGCGTATGTCCTGATAGCGGTCCGCCACACCCAGCGCGCCACAGTCGATGAATCCAGAA 2639
QY 4326 AAGCGGCCATTTCCACCATGATATTCGGCAAGCAGGATCGCCATGGTTCACGAGAGA 4385
Db 2640 AAGCGGCCATTTCCACCATGATATTCGGCAAGCAGGATCGCCATGGTTCACGAGAGA 2699
QY 4386 TCCTCGCGCTCGGCGCATCGCGCTTCGAGCTGGCGAAGCTTCGGCTGGCGCGAGCCCC 4445
Db 2700 TCCTCGCGCTCGGCGCATCGCGCTTCGAGCTGGCGAAGCTTCGGCTGGCGCGAGCCCC 2759
QY 4446 TGATGCTCTTCGTCAGATCATCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCT 4505
Db 2760 TGATGCTCTTCGTCAGATCATCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCT 2819
QY 4506 CGCTCGATGCGATGTTTCGCTTGGTTCGATGCGAGTAGGAGTACCGGATCAAGCGTATGC 4565
Db 2820 CGCTCGATGCGATGTTTCGCTTGGTTCGATGCGAGTAGGAGTACCGGATCAAGCGTATGC 2879
QY 4566 AGCGCGCGCATTCATCAGCATGATGATGATCTTCTCGGCGAGGAGGAGTGAATGAC 4625
Db 2880 AGCGCGCGCATTCATCAGCATGATGATGATCTTCTCGGCGAGGAGGAGTGAATGAC 2939
QY 4626 AGGAGATCTGCGCGCGCATTCGCGCAATAGCAGCAGTTCGCTTCCCGCTTCAGTGACA 4685
Db 2940 AGGAGATCTGCGCGCGCATTCGCGCAATAGCAGCAGTTCGCTTCCCGCTTCAGTGACA 2999
QY 4686 ACCTCGAGCACAGCTGCGCAAGAAACCGCGCTGTCGCGCAGCAGATACCGCGCTGCC 4745
Db 3000 ACCTCGAGCACAGCTGCGCAAGAAACCGCGCTGTCGCGCAGCAGATACCGCGCTGCC 3059
QY 4746 TCCTCTGCAAGTTCATTCAGGGCACCGGACAGGTGCGGTCTTGACAAAAGAACCGGGCGC 4805
Db 3060 TCCTCTGCAAGTTCATTCAGGGCACCGGACAGGTGCGGTCTTGACAAAAGAACCGGGCGC 3119
QY 4806 CCTGCGCTGACAGCGGACACGGCGCATCAGAGCAGCGATGCTGCTGTTGTCGCCAG 4865
Db 3120 CCTGCGCTGACAGCGGACACGGCGCATCAGAGCAGCGATGCTGCTGTTGTCGCCAG 3179
QY 4866 TCATAGCCGAATAGCTCTCCACCCAGCGCGCGAGAACCTGCGTGAATCCATCTTGT 4925
Db 3180 TCATAGCCGAATAGCTCTCCACCCAGCGCGCGAGAACCTGCGTGAATCCATCTTGT 3239
QY 4926 TCATATCGGAACAGCATCTCATCTGCTCTCTTGATCAGATCTTGATCCCTGCGCCAT 4985
Db 3240 TCATATCGGAACAGCATCTCATCTGCTCTCTTGATCAGATCTTGATCCCTGCGCCAT 3299
QY 4986 CAGATCTTGGCGGCAAGAAAGCCATCCAGTTTACTTTTCAGGCTTCCCAACCTTACCA 5045
Db 3300 CAGATCTTGGCGGCAAGAAAGCCATCCAGTTTACTTTTCAGGCTTCCCAACCTTACCA 3359
QY 5046 GAGGCGCCCGCAGCTGGCAATTCGGGTTTCGCTTGTGTCATTAACCGCCCGAGTACG 5105
Db 3360 GAGGCGCCCGCAGCTGGCAATTCGGGTTTCGCTTGTGTCATTAACCGCCCGAGTACG 3419

QY 5106 TATCGCCATGTAAGCCCACTGCAAGCTACCTGCTTTCTCTT 5146
Db 3420 AACTGTTGGGAAGGCGATCGTGGGCGCTTTCGCTATT 3460

RESULT 12
US-10-021-403A-9
; Sequence 9, Application US/10021403A
; Publication No. US20030074679A1
; GENERAL INFORMATION:
; APPLICANT: Adviseys
; TITLE OF INVENTION: Administration of Nucleic Acid Sequence to Female Animal to Enhance Fertility
; TITLE OF INVENTION: Growth in Offspring
; FILE REFERENCE: HO-P02021US1/100021476/OTA 00-91
; CURRENT APPLICATION NUMBER: US/10/021,403A
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/255,021
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 9
; LENGTH: 3534
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: This is a plasmid pSPC5-12-HVGRH utilized in the present invention
; OTHER INFORMATION: ON.
US-10-021-403A-9

Query Match 30.4%; Score 1603.8; DB 15; Length 3534;
Best Local Similarity 82.8%; Pred. No. 0;
Matches 2022; Conservative 0; Mismatches 27; Indels 392; Gaps 1;

QY 2706 CTAGACGTAATCATGTCATAGCTGTTCTCTGTGTAATGTTATCGGCTCACAATTC 2765
Db 1412 CTTCGGCTTAATCATGTCATAGCTGTTCTCTGTGTAATGTTATCGGCTCACAATTC 1471
QY 2766 ACACAAATCATGAGCCGGAAGCATAAAGTGTAAAGCTGGGCTGCTTAATGAGTGAGCTA 2825
Db 1472 ACACAAATCATGAGCCGGAAGCATAAAGTGTAAAGCTGGGCTGCTTAATGAGTGAGCTA 1531
QY 2826 ACTCACAATTAATGCTGTTGCGTCACTGCGCGCTTTCCAGTCGGGAAACCTGTCGTGCCA 2885
Db 1532 ACTCACAATTAATGCTGTTGCGTCACTGCGCGCTTTCCAGTCGGGAAACCTGTCGTGCCA 1591
QY 2886 GCTCATTATTAATGATCGGCAACGCGCGGAGAGCGGTTTCGCTATTCGGCGCTCTTC 2945
Db 1592 GCTCATTATTAATGATCGGCAACGCGCGGAGAGCGGTTTCGCTATTCGGCGCTCTTC 1651
QY 2946 CGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGCTGCGGAGCGGTATCAGC 3005
Db 1652 CGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGCTGCGGAGCGGTATCAGC 1711
QY 3006 TCATCAAGGCGGTAAATAGGTTATCCACAGATCAGGGGATAACGCGAGGAAGACAT 3065
Db 1712 TCATCAAGGCGGTAAATAGGTTATCCACAGATCAGGGGATAACGCGAGGAAGACAT 1771
QY 3066 GTGAGCAAAAGGCGCAGCAAAAGGCGCAGGAACCGTAAAGAGCCCGCTTTCGCGCTTTT 3125
Db 1772 GTGAGCAAAAGGCGCAGCAAAAGGCGCAGGAACCGTAAAGAGCCCGCTTTCGCGCTTTT 1831
QY 3126 CCATAGGCTCGGCGCGCTGACAGGATCAAGAGATCAAGAGATCAAGAGAGGAGGCG 3185
Db 1832 CCATAGGCTCGGCGCGCTGACAGGATCAAGAGATCAAGAGATCAAGAGAGGAGGCG 1891
QY 3186 AAACCCAGCAGGACTATAAGATACAGAGGTTTCCCGCTGGAAGCTCCCTCGTGGGCTC 3245
Db 1892 AAACCCAGCAGGACTATAAGATACAGAGGTTTCCCGCTGGAAGCTCCCTCGTGGGCTC 1951
QY 3246 TCCTGTTTCGAGCTCGCGCTTACCGGATACCTGTCGCGCTTTCCTCTTCCTTCGCGGAAGCGT 3305
Db 1952 TCCTGTTTCGAGCTCGCGCTTACCGGATACCTGTCGCGCTTTCCTCTTCCTTCGCGGAAGCGT 2011

Db	1708	TCGTCCTCGAGTTCAATTCAGGGCACCGGACAGGTCGGTCTTGACAAAAGAACCGGGCGC	1767
QY	4806	CCCTCGCTGACAGCCCGGAACAGCGGGGCACTCAGACAGCCGATGTCTGTGTGGCCGAG	4865
Db	1768	CCCTCGCTGACAGCCCGGAACAGCGGGGCACTCAGACAGCCGATGTCTGTGTGGCCGAG	1827
QY	4866	TCATAGCCGATAGCTCTCCACCACAGCGGCCGGAGAACCTGCGTGCAATCCATCTGT	4925
Db	1828	TATAGCCGATAGCTCTCCACCACAGCGGCCGGAGAACCTGCGTGCAATCCATCTGT	1887
QY	4926	TCAATCATGCGAAACGATCCTCATCCTGTCTCTTGATCAGATCTTGATCCCTCGGCCAT	4985
Db	1888	TCAATCATGCGAAACGATCCTCATCCTGTCTCTTGATCAGATCTTGATCCCTCGGCCAT	1947
QY	4986	CAGATCCTTGGCGCAGAAAGCCATCCAGTTTACTTTGCAGGGCTTCCGCACTTACCA	5045
Db	1948	CAGATCCTTGGCGCAGAAAGCCATCCAGTTTACTTTGCAGGGCTTCCGCACTTACCA	2007
QY	5046	GAGGCGGCCCCAGCTGGCAATTCCGGTTTCGCTTGCTGCTCATAAACCGCCGAGTCTAGC	5105
Db	2008	GAGGCGGCCCCAGCTGGCAATTCCGGTTTCGCTTGCTGCTCATAAACCGCCGAGTCTAGC	2067
QY	5106	TA 5107	
Db	2068	AA 2069	

RESULT 14

US-09-759-960-7/c
; Sequence 7, Application US/09759960
; Patent No. US20010006639A1

; GENERAL INFORMATION:

APPLICANT: Urban, Robert G.

APPLICANT: Chiciz, Roman M.

APPLICANT: Collins, Edward J.

APPLICANT: Hedley, Mary Lynn

TITLE OF INVENTION: IMMUNOGENIC PEPTIDES FROM THE HPV E7

INVENTION:	TITLE OF INVENTION:
PROTEIN:	PROTEIN:

NUMBER OF SEQUENCES: 33

NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:

CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson

STREET: 225 Franklin street
ADDRESS: FISH & RICHARDSON

STREET: 223 FRANKLIN STREET
CITY: Boston

CITY: BOSTON
STATE: MASTATE: MA
COUNTRY: US

COUNTRY: US
ZTP: 02110-2804

ZIP: 02110-2804
COMPUTER READABLE FORM.

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette

```

; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible

```

; IBM COMPATIBLE
; OPERATING SYSTEM: WINDOWS95

```

```

; OPERATING SYSTEM: WINDOWS95
; SOFTWARE: FASTSEC for Windows

```

```

; SOFTWARE: FASTSEQ FOR WINDOW
: CURRENT APPLICATION DATA:

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER. US/08/7

```

APPLICATION NUMBER: US/09/77
FILING DATE:

; FILING DATE:
 ; PRIOR APPLICATION DATA:

;; PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/1168

APPLICATION NUMBER: 09/169,
FILING DATE:

;; FILING DATE: 11/11/2009
; ATTORNEY / AGENT INFORMATION:

ATTORNEY/AGENT INFORMATION:

NAME: Fraser, Janis K.

REGISTRATION NUMBER: 34,819

REFERENCE/DOCKET NUMBER: 08

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-542-5070

TELEFAX: 617-543-8906

TELEX: 200154

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

```
; LENGTH: 4665 base pairs
```

; TYPE: nucleic acid

; STRANDEDNESS: double


```
Db 3695 GAAAAAGATTGGTAGTCTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGTTT 3636
Qy TTGTTTCAAGCAGCAGATTACGGCAGAAAAAAGGATCTCAAGAAGATCTCTTTGATCT 3727
Db TTGTTTCAAGCAGCAGATTACGGCAGAAAAAAGGATCTCAAGAAGATCTCTTTGATCT 3576
Qy TTCTACCGGGTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGGATTTTGGTCATGA 3787
Db TTCTACCGGGTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGGATTTTGGTCATGA 3516
Qy GATTATCGTCGACCAAGCGCCATCGTGCT----- 3819
Db GATTATCAAAAAGGATCTTACCTAGATCTTTTAAATTAAATGAAGTTTAAATCAA 3456
Qy -----CCCACTCTGCACTTCGGG 3840
Db TCTAAAGTATATATAGTAACCTGAGGCTATGGCAGGGCTGCGGCCGACGTTGGCTG 3396
Qy GCATGGATGCGCGATAGCGCTGCTGTTCTTGGATGCGGACGGATTTGCACTGCCG 3900
Db CGAGCCCTGGCCCTTCCCGAACTTGGGGGTGGGGTGGGAAAGAAAGAACCGGG 3336
Qy -----TAGAACTCCCGAGGTCGTCAGCCTCAGGCAAGCAGCTCAACCAAC 3946
Db CGTATTGCCCCAATGGGTCTCGTGGGTATCGACAGATGCCAGCCCTGGGACGAA 3276
Qy TCGGAGGGGATCGA----- 3961
Db CCCCCTGTTTATGAACAAACGACCAACCGTGGTTTTATTCTGCTTTTATTGCCG 3216
Qy -----GCCCG 3967
Db TCATAGCGCGGTTCCTTCGGTATTGTCTCTTCGTGTTTCAGTTAGCTCCCTCTAG 3156
Qy GGTGGGGAAGAAATCCAGCATAGATCCCGGCTCGAGGATCATCCAGCGCGTCC 4027
Db GGTGGGGAAGAAATCCAGCATAGATCCCGGCTCGAGGATCATCCAGCGCGTCC 3036
Qy GGAAGACGATTCGAAGCCAACTTTCATAGAGCGCGGTGGATCGAAATCTCGTG 4087
Db GGAAGACGATTCGAAGCCAACTTTCATAGAGCGCGGTGGATCGAAATCTCGTG 3036
Qy ATGCGAGGTTGGGCTCGCTGGTCAATTTGAAACCCAGAGTCCCGCTCAGAGAA 4147
Db ATGCGAGGTTGGGCTCGCTGGTCAATTTGAAACCCAGAGTCCCGCTCAGAGAA 2976
Qy CTCGTAAGAGGCGATAGAAAGGCGATCGCTGCGAATCGGAGCGGCGATACCGTAAG 4207
Db CTCGTAAGAGGCGATAGAAAGGCGATCGCTGCGAATCGGAGCGGCGATACCGTAAG 2916
Qy CACGAGGAGGCGTCAGCCCATTCGCGCAAGCTCTTCAGCAATATCACGGGTAGCCAA 4267
Db CACGAGGAGGCGTCAGCCCATTCGCGCAAGCTCTTCAGCAATATCACGGGTAGCCAA 2856
Qy CGCTATGTCCTGATAGCGGTCCGCCACACCCAGCGGCCACAGTCGATGAATCCAGAAA 4327
Db CGCTATGTCCTGATAGCGGTCCGCCACACCCAGCGGCCACAGTCGATGAATCCAGAAA 2796
Qy GCGGCCATTTCCACCATGATTTTCGCGNAGCAGGATCGCCATGGGTACAGCGGATC 4387
Db GCGGCCATTTTCACCATGATTTTCGCGNAGCAGGATCGCCATGGGTACAGCGGATC 2736
Qy CTCGCGTCGGGCTAGCGGCTTTAGCCTGGCGAAACAGTTTCGGCTGGCGGAGCCCTG 4447
Db CTCGCGTCGGGCTAGCGGCTTTAGCCTGGCGAAACAGTTTCGGCTGGCGGAGCCCTG 2676
Qy ATGCTCTTCGTCAGATCATCTGATCGAAGACCGGCTTCCATCCGATACGTGCTCG 4507
Db ATGCTCTTCGTCAGATCATCTGATCGAAGACCGGCTTCCATCCGATACGTGCTCG 2616
Qy CTCGATGCGATGTTTCGTTGGTTCGATGGCAGGTAGCGGATCAAGCGGTATGCGAG 4567
```

```
Db 2615 CTCGATGCGATGTTTCGTTGGTGGTGAATGGCAGGTAGCCGATCAAGCGTATGCGAG 2556
Qy CCGCGCATTTGCATCAGCCATGATGATATCTTTCTCGGCAAGGAGCAAGGTGAGATGACAG 4627
Db CCGCGCATTTGCATCAGCCATGATGATATCTTTCTCGGCAAGGAGCAAGGTGAGATGACAG 2496
Qy GAGATCTTGGCCCCGGCACTTCCGCCAATAGCAGCAGTCCCTTCCCGCTTCAGTGACAAC 4687
Db GAGATCTTGGCCCCGGCACTTCCGCCAATAGCAGCAGTCCCTTCCCGCTTCAGTGACAAC 2436
Qy GTGAGGCAAGCTGCGCAAGGAAGCGCCGTCGTGGCCAGCCACGATAGCCGCGTGCCTC 4747
Db GTGAGGCAAGCTGCGCAAGGAAGCGCCGTCGTGGCCAGCCACGATAGCCGCGTGCCTC 2376
Qy GTCTGCGAGTTCAATTCAGGGGACCGGACAGGTCGCTTGTGACAAAAGAACCGGGGCGCC 4807
Db GTCTTGCAGTTCAATTCAGGGGACCGGACAGGTCGCTTGTGACAAAAGAACCGGGGCGCC 2316
Qy CTGGCTGACAGCGGGAACACGGCGCATCAGAGCAGCCGATTTGTGTTGTGCCAGTC 4867
Db CTGGCTGACAGCGGGAACACGGCGCATCAGAGCAGCCGATTTGTGTTGTGCCAGTC 2256
Qy ATAGCCGAATAGCCTTCCACCCAGCGGCGGAGAACCTGCGTGCAATCCATTTGTTTC 4927
Db ATAGCCGAATAGCCTTCCACCCAGCGGCGGAGAACCTGCGTGCAATCCATTTGTTTC 2196
Qy AATCATGCGAAACGATCTCATCTGTCTCTTGATCAGATCTTG 4971
Db AATCATGCGAAACGATCTCATCTGTCTCTTGATCAGATCTTTG 2152
```

Search completed: February 15, 2004, 10:50:34

Job time : 1750 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 09:54:35 ; Search time 9963 Seconds
(without alignments)
12887.739 Million cell updates/sec

Title: US-09-921-143-36

Perfect score: 5283

Sequence: 1 aacttgacctatgcgact.....tgagtgcttgaggcagcgtg 5283

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST.*

1: em_estba:*
2: em_esthm:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rpd:*
26: em_gss_png:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	991	18.8	2450 11	AK047844
2	948	17.9	1067 9	AU081137
3	919.8	17.4	1089 9	AU081124
4	875	16.6	925 14	CB686151

5	809.2	15.3	966	29	BZ570738	BZ570738	msb2_1513
6	781.2	14.8	1163	9	AU081044	AU081044	
7	781.2	14.8	1249	29	BZ572284	BZ572284	msb2_2572
8	780.8	14.8	831	10	BG680919	BG680919	602628716
9	780.8	14.8	858	14	CA488579	CA488579	AGENCOURT
10	780.6	14.8	833	14	CB686421	CB686421	Bn01b_04j
11	774	14.7	954	9	AL044364	AL044364	DKFZp343c
12	754	14.3	1369	29	BZ579291	BZ579291	msb2_6285
13	746	14.1	844	29	BZ574513	BZ574513	msb2_3706
14	742.2	14.0	1346	29	BZ572730	BZ572730	msb2_2771
15	740.8	14.0	914	29	BZ569398	BZ569398	pacs2_164
16	740.4	14.0	831	29	TBR548748	TBR548748	Trypanoso
17	740	14.0	1194	29	BZ578985	BZ578985	msb2_6095
18	734.8	13.9	811	29	ATH517156	ATH517156	Arabidops
19	734.2	13.9	779	9	AL875347	AL875347	AL875347
20	725.6	13.7	959	29	BZ557985	BZ557985	pacs1_60
21	724	13.7	724	9	AL645114	AL645114	AL645114
22	719.6	13.6	847	28	AZ687169	AZ687169	ENTMPO5TF
23	719.6	13.6	854	29	BZ570648	BZ570648	msb2_1469
24	713.4	13.6	896	29	BZ551301	BZ551301	pacs1_60
25	713.4	13.6	800	9	AJ281449	AJ281449	4A3A-P4D5
26	715.8	13.5	956	9	AL959357	AL959357	AL959357
27	713.6	13.5	921	11	CNS08VR4	CNS08VR4	Single re
28	711	13.5	1004	9	AJ281480	AJ281480	4A3A-P4G8
29	710.2	13.4	715	9	AL661706	AL661706	AL661706
30	708	13.4	828	29	BZ573391	BZ573391	msb2_3083
31	707.6	13.4	835	29	BZ571572	BZ571572	msb2_1950
32	707	13.4	707	9	AL658688	AL658688	AL658688
33	704	13.3	818	12	B1180988	B1180988	UNL-P-F2-
34	703.4	13.3	705	9	AL635845	AL635845	AL635845
35	703.4	13.3	705	9	AL643164	AL643164	AL643164
36	700.2	13.3	838	28	AZ679307	AZ679307	ENTLGR4TF
37	700	13.3	700	13	BQ660293	BQ660293	HT01N05W
38	699.6	13.2	718	9	AL631067	AL631067	AL631067
39	699	13.2	699	13	BQ659852	BQ659852	HF01O22W
40	697.2	13.2	894	29	BZ578323	BZ578323	msb2_5785
41	695.8	13.2	700	13	BQ660226	BQ660226	HT01J04W
42	695	13.2	695	13	BQ660140	BQ660140	HT01G11W
43	695	13.2	697	13	BQ660197	BQ660197	HT01G17W
44	695	13.2	697	13	BQ660261	BQ660261	HT01L03W
45	695	13.2	700	14	CA305674	CA305674	EST000022

ALIGNMENTS

RESULT 1
AK047844
LOCUS
DEFINITION
AK047844 2450 bp mRNA linear HTC 05-DEC-2002
Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched library, clone:C130004M09 product:vascular endothelial growth factor C, full insert sequence.

ACCESSION
AK047844
VERSION
AK047844.1 GI:26339001
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
MEDLINE
PUBMED 10349636

REFERENCE
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
MEDLINE
PUBMED 11042159

REFERENCE
AUTHORS

3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Kono,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kaehiagi,K.,
Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Wataniki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multiplexed capillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

20530913
11076861
4 Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
Arakawa,T., Hara,A., Fukunishi,Y., Kono,H., Adachi,J., Fukuda,S.,
Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamakawa,I.,
Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
Fleischmann,W., Gaasterland,T., Giusti,C., King,B., Kochiwa,H.,
Kuehl,P., Lewis,S., Matsumoto,Y., Nikaide,T., Pesole,G.,
Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
Marchionni,L., Mashima,J., Mazzairelli,J., Mombaerts,P., Nordone,P.,
Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
and Hayashizaki,Y.

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851
5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2450)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,P., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,I., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/

FEATURES

Location/Qualifiers

source

1. .2450
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:Cl30004M09"
/db_xref="taxon:10090"
/clone="Cl30004M09"
/tissue_type="head"
/dev_stage="16 days embryo"
/note="unnamed protein product; putative
vascular endothelial growth factor C (MGD|GI:109124,
GB|NM_009506, evidence: BLASTN, 99%, match=1784)"
/codon_start=1
/protein_id="BAC33172.1"
/db_xref="GI:26339002"
/translation="WHILCFSLACLSLAALIPSPREAPATVAAPESGLPSEAEPPD
GGEVKAPEKDELEQLRSVSDVDELSVLIPYWMYKCLAKGKQKPTLNRIGDS
VFAAHAYNTEILKSIDNWRATQCMPREVIDVGKFGAANTFFKPCVSVYRCGG
CNSEGLQCMNTSTGLSKTLFEITVPLSQGPKPTVIFANHTSCRMKSLDVYQVH
SIIRSLPATLPQQAANKTPTVNMNRCMLAQDQFIYSNVEDDSTNGFDVFC
GNKLEIDETCCVCKGILRPSGPHKELDRDSCQVCKNKLFPNSCANREFENT
COCVCKTICPRNQLNPGKCAECENTQKCFKKGKHFHQTCSCVRRPCANRLKHD
PSLSSEEVCRVPSVWKEPHLN"

CDS

BASE COUNT 714 a 541 c 558 g 637 t
ORIGIN
Query Match 18.8%; Score 991; DB 11; Length 2450;
Best Local Similarity 85.4%; Pred. No. 6.1e-266;
Matches 1121; Conservative 0; Mismatches 180; Indels 12; Gaps 1;
QY 481 CACCTGGTGGATCGATCCATCATGCACTCGTGGGCTTCTCTCTGGCGGTGTTCT 540
Db 151 CTTCCCTGCTCCCGGTCCATCCACATGCTGCTTGTGCTTGTCTCTGGCGGTGTTCC 210
QY 541 CTGCTCGCGCTCGCGTCTCCCGGTCTCTCGAGAGCGCCCGCGCGCGCGCGCTTC 600
Db 211 CTGCTCGCGCTCGCGTCTGATCCCGAGTCCGCGAGGCGCCCGCGCGCGCTTC 270
QY 601 GAGTCCGAGCTCGACTCTTCGAGCGGAGCGCGCGCGCGGTGAGGCCACGGCTTATGCA 660
Db 271 GAGTCCGAGCTCGCGTCTTCGGAAGCGGAGCGCGCGCGCGGTGAGGCCACGGCTT 330
QY 661 AGCAAGAGTCTGAGGAGCAGTTACGCTCTGTCAGCTAGTAGTGAATCTAGTCTGTA 720
Db 331 GGCAAGAGCTGAGGAGCAGTTGCGCTCTGTCAGCTAGTAGTGAATCTAGTCTGTC 390
QY 721 CTCTACCCAGATATTGGAAAATGTACAAGTCTCAGCTAAGGAAAAGAGGCTGGCAACAT 780
Db 391 CTGTACCCAGACTACTTGGAAAATGTACAAGTCCAGCTCGGAAAAGGCGGTGGC---- 445
QY 781 AACAGAGACAGGCCAACCTCAACTCAAGGACAGAGAGACTATATAAATTTGCTGCAGCA 840
Db 446 -----AGCAGCCCAACCTCAATACAGGACAGGAGTGTAAAAATTTGCTGCTGCA 498
QY 841 CATTATAACACAGAGATCTTGAAAAGTATTGATAATAGTGAGGAGAAAGACTCAATGCATG 900
Db 499 CATTATAACACAGAGATCTTGAAAAGTATTGATAATAGTGAGGAGAAAGACTCAATGCATG 558
QY 901 CCACGGAGGTGTGATAGATGTGGGAGAGAGTTTGGAGTCCGACAAAACACCTTCTTT 960
Db 559 CCACGTGAGGTGTGATAGATGTGGGAGAGAGTTTGGAGCAGCCAGCAAAACACCTTCTTT 618
QY 961 AAACCTCCATGTGTGCTCTACAGATGTGGGGTGTCTGCAATAGTGAGGGGTGCGAG 1020
Db 619 AAACCTCCATGTGTGCTCTACAGATGTGGGGTGTCTGCAACAGCGAGGGGTGCGAG 678
QY 1021 TGATGAACACACAGCAGGCTACTCAGCAAGAGCTTATTGAAATTTACAGTGCCTCTC 1080
Db 679 TGATGAACACACAGCAGGTTTACCTCAGCAAGAGCTTATTGAAATTTACAGTGCCTCTC 738
QY 1081 TCTCAAGGCCCCAAACCAAGTAACATCACTTTTGCATCACTTCTCTCCGATGCATG 1140

Db	739	TCACAGGCCCCAACCAAGTCACATCAGTTTCCCAATCACATCTCTCCCGGTGATG	798
QY	1141	TCTAACTGGATGTTTACAGACAAGTTCATTCATTATTAGAGCTTCCCTGCCAGCAACA	1200
Db	799	TCTAACTGGATGTTTACAGACAAGTTCATTCATTATTAGAGCTTCTCTGCCAGCAACA	858
QY	1201	CTACACAGTGTCCAGGAGGACCAAGACAGCTGCCCAACCAATACATGTGGATATACAC	1260
Db	859	TTACACAGTGTCCAGGAGGACCAAGACAGCTGCCCAACCAATATGTGGATATACAC	918
QY	1261	ATCTGCAGATCCCTGGCTCAGGAGATTTTATGTTTTCCTCGATGCTGGAGATGACTCA	1320
Db	919	ATGTGCCGATCCCTGGCTCAGGAGATTTTATGTTTTCCTCGATGCTGGAGATGACTCA	978
QY	1321	ACAGATGATTCATGATCATCTGTGGACCAACAGAGAGCTGGATGAAGAGACTGTTCAG	1380
Db	979	ACCAATGGATTCATGATCATCTGTGGACCAACAGAGAGCTGGATGAAGAGACTGTTCAG	1038
QY	1381	TGTGTCTGCAGAGGGGCTTCGGCTGCCAGCTGTGGACCCCAACCAAGAACTAGACAGA	1440
Db	1039	TGTGTCTGCAGAGGGGCTTCGGCTGCCAGCTGTGGACCCCAACCAAGAACTAGACAGA	1098
QY	1441	AACATATGCTAGTGTGTCTGTAACCAAACTCTTCCCAAGCAATGTGGGGCAACCGA	1500
Db	1099	GACTCATGTGTGTGTCTGTAACCAAACTCTTCCCAAGCAATGTGGGGCAACCGA	1158
QY	1501	GAATTTGATGAACACATGCTGCTGTATGTAAGAGAACTGCTCCCAAGCAATGTGGGGCAACCGA	1560
Db	1159	GAATTTGATGAACACATGCTGCTGTATGTAAGAGAACTGCTCCCAAGCAATGTGGGGCAACCGA	1218
QY	1561	CTAATCTCTGAAATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTGTAATA	1620
Db	1219	CTGAATCTCTGAAATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTGTAATA	1278
QY	1621	GGAAAGATTTCCACCAACCAATGCTGCTGTATGTAAGAGAACTGCTCCCAAGCAATGTGGGGCAACCGA	1680
Db	1279	GGAAAGATTTCCACCAACCAATGCTGCTGTATGTAAGAGAACTGCTCCCAAGCAATGTGGGGCAACCGA	1338
QY	1681	AAGCTTGTGAGCCAGATTTTCATATAGTGAAGAGTGTGCTGTGTCCTTCATAT	1740
Db	1339	AAGCAATGTGATCCAGAGTGTCTTATGTAAGAGTATGCTGCTGTGTCCTTCATAT	1398
QY	1741	TGGAAGAGACCAAAATGAGCTAATCTAGGATCCGTACCTGCCAGGCTTTT	1793
Db	1399	TGGAAGAGGCCACATCTGAACATAGATCATACAGTTTTCAGTCCACCAATTTT	1451
RESULT 2			
AC081137			
LOCUS			
DEFINITION			
SOURCE			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			

FEATURES	Location/Qualifiers	
source	1..1067	
	/organism="Oncorhynchus mykiss"	
	/mol_type="mRNA"	
	/db_xref="taxon:8022"	
	/clone="K12"	
	/tissue_type="kidney"	
	/clone_lib="Oncorhynchus mykiss Kidney infected by infectious hematopoietic necrosis virus"	
	/note="common name:rainbow trout ; infected by infectious hematopoietic necrosis virus"	
BASE COUNT	235 a 292 c 288 g 250 t	2 others
ORIGIN		
	Query Match	
	Best Local Similarity	17.9% ; Score 948 ; DB 9 ; Length 1067 ;
	Matches	976 ; Conservative 0 ; Mismatches 10 ; Indels 2 ; Gaps 2 ;
QY	2706	CTAGACGTAATCATGCTCATAGCTGTTTCCCTGTGTGAATTTGTTATCCGCTCACAATTC 2765
Db	82	CTTGGCGTAATCATGCTCATAGCTGTTTCCCTGTGTGAATTTGTTATCCGCTCACAATTC 141
QY	2766	ACACAACATACAGCGGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTACGCTA 2825
Db	142	ACACAACATACAGCGGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTACGCTA 201
QY	2826	ACTCATATTAATGGCTTGGCTCACTGCCCGCTTTCAGTCCGGAACCTGTCTGTGCA 2885
Db	202	ACTCATATTAATGGCTTGGCTCACTGCCCGCTTTCAGTCCGGAACCTGTCTGTGCA 261
QY	2886	GCTGCATTAAATGAATCGGCCCAACCGCGGGAGAGCGGTTTGGCTATTGGCGCTCTTC 2945
Db	262	GCTGCATTAAATGAATCGGCCCAACCGCGGGAGAGCGGTTTGGCTATTGGCGCTCTTC 321
QY	2946	CGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 3005
Db	322	CGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 381
QY	3006	TCACATCAAGCGGTAATACGGTTATCACAGATCAGGGATACCGGATACCGGAGAAACAT 3065
Db	382	TCACATCAAGCGGTAATACGGTTATCACAGATCAGGGATACCGGATACCGGAGAAACAT 441
QY	3066	GTAGCAAAAGGCGCAGCAAAAGGCGGCAACCGTAAAGGCGCGTTTGTCTGGCGTTT 3125
Db	442	GTAGCAAAAGGCGCAGCAAAAGGCGGCAACCGTAAAGGCGCGTTTGTCTGGCGTTT 501
QY	3126	CCATAGGCTCGCGCGCGCTGACAGCATCAAAAATCAGCTCAAGTCAAGTCAAGTCA 3185
Db	502	CCATAGGCTCGCGCGCGCTGACAGCATCAAAAATCAGCTCAAGTCAAGTCAAGTCA 561
QY	3186	AAACCCGACAGACTATATAAGATACAGGCGGTTTCCCGCTGGAAGCTCCCTCGTGGCTC 3245
Db	562	AAACCCGACAGACTATATAAGATACAGGCGGTTTCCCGCTGGAAGCTCCCTCGTGGCTC 621
QY	3246	TCCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCCTCGTGGGAAAGCT 3305
Db	622	TCCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCCTCGTGGGAAAGCT 681
QY	3306	GGCGGTTTCTCATAGCTCACGCTGATGATCTCAGTTCGCTGAGTTCGCTTCCTCCAA 3365
Db	682	GGCGGTTTCTCATAGCTCACGCTGATGATCTCAGTTCGCTGAGTTCGCTTCCTCCAA 741
QY	3366	GCTGGGCTGTGTGACCAAGAACCCCGCTTACCGCGAGCGCTGCGCTTATCCGCTAATA 3425
Db	742	GCTGGGCTGTGTGACCAAGAACCCCGCTTACCGCGAGCGCTGCGCTTATCCGCTAATA 801
QY	3426	TGCTTTGAGTCCCAACCGGTAAGACAGACTTATCGCACTGGCAGGACCACTGGTAA 3485
Db	802	TGCTTTGAGTCCCAACCGGTAAGACAGACTTATCGCACTGGCAGGACCACTGGTAA 861
QY	3486	CAGGATTAAGCAGGAGGATGTAGGCGGTGCTACAGAGTTCTTTGAAGTGGTGGCTTAA 3545
Db	862	CAGGATTAAGCAGGAGGATGTAGGCGGTGCTACAGAGTTCTTTGAAGTGGTGGCTTAA 921

QY	3546	CTACGGCTACACTAGAACACAGATATTTGGTATCTGCGCTCTGCTGTAAGCCAGTTACCTT	3605
Db	922	CTACGGCTACACTAGAACACAGATATTTGGTATCTGCGCTCTGCTGTAAGCCAGTTACCTT	981
QY	3606	CGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCGCTGCTAGCGTGGTTT	3665
Db	982	CGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAA-CAACGCTGGTAGCGTGT	1040
QY	3666	TTTGTGTTGAACAGCAGATTTACGGC	3693
Db	1041	TTTGTGTTGC-AGCAGCAGATTTACGGC	1067
RESULT 3			
LOCUS	AU081124	1089 bp mRNA linear	EST 30-JUL-2002
DEFINITION	AU081124	Oncorhynchus mykiss Kidney infected by infectious hematopoietic necrosis virus	Oncorhynchus mykiss cDNA clone KG'12, mRNA sequence.
ACCESSION	AU081124		
VERSION	AU081124.1	GI:6431472	
KEYWORDS			
SOURCE	Oncorhynchus mykiss (rainbow trout)		
ORGANISM	Oncorhynchus mykiss		
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.		
TITLE	1. (bases 1 to 1089)		
JOURNAL	Kono,T., Sakai,M. and LaPatra,S.E.		
COMMENT	Expressed Sequence tag Analysis of Kidney and Gill Tissues from Rainbow Trout (Oncorhynchus mykiss) Infected with Infectious Hematopoietic Necrosis Virus		
Mar. Biotechnol.	2 (5), 493-498 (2001)		
Contact: Masahiro Sakai			
Faculty of Agriculture			
Miyazaki University			
1-1 nishi gakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan			
Email: m.sakai@cc.miyazaki-u.ac.jp.			
Location/Qualifiers			
source			
1. .1089			
/organism="Oncorhynchus mykiss"			
/mol_type="mRNA"			
/db_xref="taxon:8022"			
/clone="KG'12"			
/tissue_type="kidney"			
/clone_lib="Oncorhynchus mykiss Kidney infected by infectious hematopoietic necrosis virus"			
/note="common name:rainbow trout ; infected by infectious hematopoietic necrosis virus"			
BASE COUNT	245 a	297 c	295 g
ORIGIN			
Query Match	17.4%;	Score 919.8;	DB 9;
Best Local Similarity	97.2%;	Pred. No. 3.5e-246;	
Matches	979;	Conservative	0; Mismatches 22; Indels 6; Gaps 4;
QY	2706	CTAGAGTAATCATGTCTACAGTCTTTCCTGTGTGAATTTGTTATCCGCTCACAATTC	2765
Db	83	CTTGGCGTAATCATGTCTACAGTCTTTCCTGTGTGAATTTGTTATCCGCTCACAATTC	142
QY	2766	ACACACATACGACCGGAGCATAAAGTGAACGCTGGGTGCTTAATGATGAGCTA	2825
Db	143	ACACACATACGACCGGAGCATAAAGTGAACGCTGGGTGCTTAATGATGAGCTA	202
QY	2826	ACTCACATTAATTTGGTTCGCTACTGCCGCTTCCAGTCGGGAAACCTGTGTCGCCA	2885
Db	203	ACTCACATTAATTTGGTTCGCTACTGCCGCTTCCAGTCGGGAAACCTGTGTCGCCA	262
QY	2886	GCTGCATTAATGATCGGCACACGCGGGAGAGCGGTTTGGTATTTGGGCGCTCTTC	2945
Db	263	GCTGCATTAATGATCGGCACACGCGGGAGAGCGGTTTGGTATTTGGGCGCTCTTC	322

JOURNAL
COMMENT

Unpublished
Contact: Singh, J.A.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-food Canada
KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A
0C6, Canada
Tel: (613) 759-1662
Fax: (613) 759-1701
Email: singhja@em.agr.ca.
Location/Qualifiers
source
1. .925

FEATURES

/organism="Brassica napus"
/mol_type="mRNA"
/cultivar="Westar"
/db_xref="taxon:3708"
/clone="Bn01b_02c08"
/tissue_type="fourth leaf"
/dev_stage="3 weeks seedling grown at room temperature"
/clone_lib="Bn01b_AAFRC_ECORC transgenic Brassica napus_ove
repressing_ENCBF17 constitutively frost_tolerant"
/notes="Vector: Bluescript SK+/XhoI-EcoRI; Site_1: EcoRI;
Site_2: XhoI; Germinated in soil flats and seedlings grown
for 3 weeks in a Conviron E-15 cabinet set at 20°C / 16 hr
light (250 Em-2sec-1) and 16 °C / 8 hr dark. Fourth leaves
collected at 9 am and immediately frozen."
BASE COUNT 218 a 245 c 253 g 203 t
ORIGIN

Query Match 16.6%; Score 875; DB 14; Length 925;
Best Local Similarity 98.6%; Pred. No. 1.2e-233;
Matches 886; Conservative 6; Mismatches 6; Indels 1; Gaps 1;

QY 2701 GCGTGTAGACGTAATCATGCTCATAGCTGTTTCCTGTGTGAATGTTATCGGTCACA 2760
DB GGCGGTGGCGTAATCATGCTCATAGCTGTTTCCTGTGTGAATGTTATCGGTCACA 840
QY 2761 ATTCCACACATACGAGCGGAAGCATAAAGTGTAAGCCTGGGGTCCCTAATGAGTG 2820
DB ATTCCACACATACGAGCGGAAGCATAAAGTGTAAGCCTGGGGTCCCTAATGAGTG 780
QY 2821 AGTAATCATTAATGGTTCGCTCACTGCCCGCTTCAGTCGGGAACCTGTG 2880
DB AGTAATCATTAATGGTTCGCTCACTGCCCGCTTCAGTCGGGAACCTGTG 720
QY 2881 TGCCAGCTGATTAATGAATCGGCCAACGCGCGGAGAGCGGTTTGGTATTGGGCGC 2940
DB TGCCAGCTGATTAATGAATCGGCCAACGCGCGGAGAGCGGTTTGGTATTGGGCGC 660
QY 2941 TCTTCGGTTCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 3000
DB TCTTCGGTTCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 600
QY 3001 TCAGCTCACTCAAGGCGGTAAATACGTTATCCACAGATCAGGGATACCGCAGGAAG 3060
DB TCAGCTCACTCAAGGCGGTAAATACGTTATCCACAGATCAGGGATACCGCAGGAAG 540
QY 3061 AACATGTGACAAAGCCAGCAAAAGCGGAGAACCGGTAAAGAGCCCGCTTGTGGCG 3120
DB AACATGTGACAAAGCGGAGCAAAAGCGGAGAACCGGTAAAGAGCCCGCTTGTGGCG 480
QY 3121 TTTTTCATAGGTCGCGCCCTGACGAGCATCACAATAATCGAGCTCAAGTCAGAGG 3180
DB TTTTTCATAGGTCGCGCCCTGACGAGCATCACAATAATCGAGCTCAAGTCAGAGG 420
QY 3181 TGGCGAAACCCGACAGACTATAAAGATACAGCGGTTTCCCTCGTGAAGCTCCCTCGTG 3240
DB TGGCGAAACCCGACAGACTATAAAGATACAGCGGTTTCCCTCGTGAAGCTCCCTCGTG 360
QY 3241 CGCTCTCTGTCGACCCCTGCGCTTACCGGATACCTGTGCGCTTTCTCCCTTCGGA 3300
DB CGCTCTCTGTCGACCCCTGCGCTTACCGGATACCTGTGCGCTTTCTCCCTTCGGA 300
QY 3301 AGCGTGGCGCTTCTCATAGCTCACGCTGTAGTATCTCAGTTCGTTGAGTTCGTCG 3360

Db 299 AGCGTGGCGCTTCTCATAGCTCAAGCTGATAGTATCTCAGTTCGTTAGTTCGTCGC 240
QY 3361 TCCAGCTGGGCTGTGTGTCAGAACCCCGCTTCAGCCCGACCGCTCGCTTATCCGCT 3420
Db 239 TCCAGCTGGGCTGTGTGTCAGAACCCCGCTTCAGCCCGACCGCTCGCTTATCCGCT 180
QY 3421 AACTATCGTCTTGTAGTCCAAACCCCGTAAAGACACAGCTTATCGCCACTGGCAGCAGCACT 3480
Db 179 AACTATCGTCTTGTAGTCCAAACCCCGTAAAGACACAGCTTATCGCCACTGGCAGCAGCACT 120
QY 3481 GGTAAACGATTAAGCAGGAGGATGTAGTGGGCTGTACAGAGTTCCTGAGTGGTGG 3540
Db 119 GGTAAACGATTAAGCAGGAGGATGTAGTGGGCTGTACAGAGTTCCTGAGTGGTGG 60
QY 3541 CCTAACTACGGCTACACTAGAGAACAGTATTTGGTATCTGCGCTCTGCTGAGCCAGT 3599
Db 59 CTAACTACGGCTACACTAGAGAACAGTATTTGGTATCTGCGCTCTGCTGAGCCAGT 1

RESULT 5
B2570738
LOCUS
DEFINITION msh2_1513.x1 msh Pseudomonas aeruginosa genomic clone msh2_1513,
genomic survey sequence.
ACCESSION B2570738
VERSION B2570738.1 GI:27205799
KEYWORDS
SOURCE
ORGANISM Pseudomonas aeruginosa
Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE
AUTHORS 1 (bases 1 to 966)
Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
Burns, J.L., Kaul, R. and Olsen, M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol., (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun
FEATURES
source
1. .966
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="MSH"
/db_xref="taxon:287"
/clone="msh2_1513"
/clone_lib="msh"
/note="Environmental isolate. Whole genomic shotgun
library."
BASE COUNT 223 a 267 c 266 g 209 t
ORIGIN
Query Match 15.3%; Score 809.2; DB 29; Length 966;
Best Local Similarity 97.0%; Pred. No. 3.7e-215;
Matches 845; Conservative 0; Mismatches 24; Indels 2; Gaps 2;
QY 2890 CATTAAATGAATCGGCCAACCGCGGAGAGCGGTTTGGTATTGGCGCTCTTCGCT 2949
Db 94 CATTAAATGAATCGGCCAACCGCGGAGAGCGGTTTGGTATTGGCGCTCTTCGCT 153
QY 2950 TCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 3009
Db 154 TCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 213
QY 3010 TCAAAGCGGTAAATACGGTTATCCAGAGATCAGGGGATAACGAGGAAGACATGTGA 3069

Db 214 TCAAGGCGGTAAATACGGTTATTCACAGAAATCAGGGATACGAGGAAAGAACATGTGA 273
 QY 3070 GCAAAAGCCAGCAAAAGCCAGAGAACCGTAAAGAGCCGGTGTGCTGGGTTTTTCCAT 3129
 Db 274 GCAAAAGCCAGCAAAAGCCAGAGAACCGTAAAGAGCCGGTGTGCTGGGTTTTTCCAT 333
 QY 3130 AGGCTCCGCCCTTGACGAGATCACAAAAATCAGCTCAAGTCAAGAGTGCGGAAAC 3189
 Db 334 AGGCTCCGCCCTTGACGAGATCACAAAAATCAGCTCAAGTCAAGAGTGCGGAAAC 393
 QY 3190 CGCAGGAGCTATAAGATACAGCGGTTTCCCTCGAAGCTCCCTCGTGGCTCTCCT 3249
 Db 394 CGCAGGAGCTATAAGATACAGCGGTTTCCCTCGAAGCTCCCTCGTGGCTCTCCT 453
 QY 3250 GTTCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCCCTTCGGGAAGCGGCG 3309
 Db 454 GTTCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCCCTTCGGGAAGCGGCG 513
 QY 3310 CTTTCTCATAGCTACGCTGAGTATCTCAGTTTCGGTGTAGGTCTGCTCCAGCTG 3369
 Db 514 CTTTCTCATAGCTACGCTGAGTATCTCAGTTTCGGTGTAGGTCTGCTCCAGCTG 573
 QY 3370 GGCTGTGTGACGAAACCCCGTTTCAGCCGACCGCTGCGCTTATCCGGTAACTATCGT 3429
 Db 574 GGCTGTGTGACGAAACCCCGTTTCAGCCGACCGCTGCGCTTATCCGGTAACTATCGT 633
 QY 3430 CTTGAGTCCAAACCGGTGAAGACAGACTTATCGGCATGCGCAGCAGCTGTGAACAGG 3489
 Db 634 CTTGAGTCCAAACCGGTGAAGACAGACTTATCGGCATGCGCAGCAGCTGTGAACAGG 693
 QY 3490 ATTACAGAGGAGTATGTAGGCGGTCTACAGAGTTCTTGAAGTGTGCGCTTAACCTAC 3549
 Db 694 NATACAGAGGAGTATGTAGGCGGTCTACAGAGTTCTTGAAGTGTGCGCTTAACCTAC 753
 QY 3550 GGCTACACTAGAAAGATTTGGTATCTGCGCTGTCTGAAAGCAGATTAACCTCGGA 3609
 Db 754 GGCTACACTAGAAAGATTTGGTATCTGCGCTGTCTGAAAGCAGATTAACCTCGGA 812
 QY 3610 AAAGAGTTGGTGTAGCTTTGATTCGGGCAAAACAAACCCGCTGTGAGCGGTGTTTTT 3669
 Db 813 AAAGAGTTGGTGTAGCTTTGATTCGGGCAAAACAAACCCGCTGTGAGCGGTGTTTTT 872
 QY 3670 GTTTCGACGAGCAGATTAACGCGCAAAAGATCTCAAGAGATTCCTTGAATCTTT 3729
 Db 873 TGTTCGAAGGCAACAAATACGCGCAAAAGATCTCAAGAGATTCCTTGAATCTTT 932
 QY 3730 TCTACGGGCTGACGCTCAGTGGAAACGAA 3760
 Db 933 T-TACGGGTTTCGACGCTCAAGGACCGAAA 962

RESULT 6

AU081044

LOCUS

DEFINITION

hematopoietic necrosis virus Oncorhynchus mykiss cdna clone KBI,
 mRNA sequence.

ACCESSION

AU081044

VERSION

AU081044.1

KEYWORDS

EST.

SOURCE

ORGANISM

Oncorhynchus mykiss

(rainbow trout)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Miyazaki University
 1-1 nishi gakuenbanadai, Miyazaki, Miyazaki 889-2192, Japan
 Email: m.sakai@cc.miyazaki-u.ac.jp.
 Location/Qualifiers
 1..1163
 /organism="Oncorhynchus mykiss"
 /mol_type="mRNA"
 /db_xref="taxon:8022"
 /clone="KBI"
 /tissue_type="kidney"
 /clone_lib="Oncorhynchus mykiss Kidney infected by
 infectious hematopoietic necrosis virus"
 /note="common name:rainbow trout ; infected by infectious
 hematopoietic necrosis virus"
 BASE COUNT 272 a 307 c 298 g 286 t
 ORIGIN

Query Match 14.8%; Score 781.2; DB 9; Length 1163;
 Best Local Similarity 99.5%; Pred. No. 3.1e-207;
 Matches 794; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 QY 2706 CTAGACGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTTGTTATCCGTCACAAATCC 2765
 Db 367 CTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTTGTTATCCGTCACAAATCC 426
 QY 2766 ACACACATACAGCGGGAAGCATAAAGTAAAGCTTGGGTGCTTAATGATGAGCTA 2825
 Db 427 ACACACATACAGCGGGAAGCATAAAGTAAAGCTTGGGTGCTTAATGATGAGCTA 486
 QY 2826 ACTCACATTAATTGGCTTTCGCTCACTCCCGCTTTCCAGTCGGGAAACCTGTCTGTGCA 2885
 Db 487 ACTCACATTAATTGGCTTTCGCTCACTCCCGCTTTCCAGTCGGGAAACCTGTCTGTGCA 546
 QY 2886 GTTCGATTAATGAATCGCCACACGCGGGGAGAGCGGTTTGGTATTTGGCGCTCTTC 2945
 Db 547 GTTCGATTAATGAATCGCCACACGCGGGGAGAGCGGTTTGGTATTTGGCGCTCTTC 606
 QY 2946 CGCTTCTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 3005
 Db 607 CGCTTCTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 666
 QY 3006 TCACCTCAAGGCGGTAAATACGTTATCCAGAAATCAGGGATACGAGGAAAGAAAT 3065
 Db 667 TCACCTCAAGGCGGTAAATACGTTATCCAGAAATCAGGGATACGAGGAAAGAAAT 726
 QY 3066 GTGAGCAAAAGCCAGCAAAAGCCAGCAAAAGCCAGCAAAAGCCAGCAAAAGCCAGCA 3125
 Db 727 GTGAGCAAAAGCCAGCAAAAGCCAGCAAAAGCCAGCAAAAGCCAGCAAAAGCCAGCA 786
 QY 3126 CCATAGGCTCCGCCCTCTGACGAGCATCAAAAATCGAAGCTCAAGTCAAGTCAAGTGGCG 3185
 Db 787 CCATAGGCTCCGCCCTCTGACGAGCATCAAAAATCGAAGCTCAAGTCAAGTCAAGTGGCG 846
 QY 3186 AAACCCGACGAGCTATAAGATACAGCGGTTTCCCTTGAAGCTCCCTCGTGGCTC 3245
 Db 847 AAACCCGACGAGCTATAAGATACAGCGGTTTCCCTTGAAGCTCCCTCGTGGCTC 906
 QY 3246 TCTGTTCGACCTCCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGT 3305
 Db 907 TCTGTTCGACCTCCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGT 966
 QY 3306 GGGCTTTCTCATAGCTCAGCTGAGTATCTCAGTTTCGGTGTAGGTGCTGCTCCAA 3365
 Db 967 GGGCTTTCTCATAGCTCAGCTGAGTATCTCAGTTTCGGTGTAGGTGCTGCTCCAA 1026
 QY 3366 GTTGGGCTGTGTGACAGAACCCCGCTTTCAGCCGACCGCTGCGCTTATTCGGTAACTA 3425
 Db 1027 GCTGGGCTGTGTGACAGAACCCCGCTTTCAGCCGACCGCTGCGCTTATTCGGTAACTA 1085
 QY 3426 TGTCTTTGAGTCAACCCCGGTGAAGACAGACTTATCGCCACTGGCAGAGCCACTGGTAA 3485
 Db 1086 TGTCTTTGAGTCAACCCCGGTGAAGACAGACTTATCGCCACTGGCAGAGCCACTGGTAA 1145

Qy	3486	CAGGATTAGCAGCGAG	3503
Db	1146	CAGGATTAGCAGCGAG	1163
RESULT 7	BZ572284	1249 bp	linear
BZ572284	msh2_2572.x3 msh Pseudomonas aeruginosa genomic clone msh2_2572,	GSS 17-DEC-2002,	
LOCUS	genomic survey sequence.		
DEFINITION	BZ572284		
ACCESSION	BZ572284		
VERSION	BZ572284.1		
KEYWORDS	GSS.		
SOURCE	Pseudomonas aeruginosa		
ORGANISM	Pseudomonas aeruginosa		
REFERENCE	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;		
AUTHORS	1 (bases 1 to 1249)		
TITLE	Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,		
JOURNAL	Burns,J.L., Kaul,R. and Olsen,M.V.		
COMMENT	Whole-Genome-Sequene variation among multiple isolates of		
	Pseudomonas aeruginosa library		
	J. Bacteriol., (2002) In press		
	Contact: Chris K. Raymond		
	Genome Center		
	University of Washington		
	Box 352145, Seattle, WA 98105-2145, USA		
	Tel: 20622216954		
	Fax: 2066857244		
	Email: craymond@u.washington.edu		
	Class: shotgun.		
FEATURES	Location/Qualifiers		
source	1..1249		
	/organism="Pseudomonas aeruginosa"		
	/mol_type="genomic DNA"		
	/strain="MSH"		
	/db_xref="taxon:287"		
	/clone="msh2_2572"		
	/clone_lib="msh"		
	/note="Environmental isolate. Whole genomic shotgun		
	library."		
BASE COUNT	292 a 337 c 323 g 294 t	3 others	
ORIGIN			
Query Match	14.8%; Score 781.2; DB 29; Length 1249;		
Best Local Similarity	95.4%; Pred. No. 3.2e-207;		
Matches	857; Conservative 0; Mismatches 35; Indels 6; Gaps 5;		
Qy	2883	CCAGCTGCATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTCGGTATTGGGCGCTC	2842
Db	100	CCAGCTGCATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTCGGTATTGGGCGCTC	159
Qy	2943	TTCCGCTTCCTCGCTCACTCGCTCGCTCGCTCGGTTCGGTGTCGGCGAGCGGTATC	3002
Db	160	TTCCGCTTCCTCGCTCACTCGCTCGCTCGGTTCGGTGTCGGCGAGCGGTATC	219
Qy	3003	AGCTCATCAAAGCGCGTAATA CGGTTATCCA CAGAATCAGGGGATAAACGAGAAAGAA	3062
Db	220	AGCTCACTAAAGCGCGTAATACGGTTATCCACAGAATCAGGGGATAAACGAGAAAGAA	279
Qy	3063	CATGTGAGAAAAGGCCAGCAAAAGGCCAGAACCGTAAAAAGCCCGTTGCTGGCGTT	3122
Db	280	CATGTGAGCAAAAGGCCAGCAAAAGGCCAGAACCGTAAAAAGCCCGTTGCTGGCGTT	339
Qy	3123	TTTCCATAGGCTCGCCCCCTCGACGAGCATCAAAAAATCGAGCGTCAAGTCAGAGGTG	3182
Db	340	TTTCCATAGGCTCGCCCCCTCGACGAGCATCAAAAAATCGAGCGTCAAGTCAGAGGTG	399
Qy	3183	GCAGAAACCGACAGACTATAAAGATACACAGCGGTTTTCCCCTGGAGCTCCCTCGTGC	3242
Db	400	GCAGAAACCGACAGACTATAAAGATACACAGCGGTTTTCCCCTGGAGCTCCCTCGTGC	459
Qy	3243	CTCTCTGTTCCGACCCTGGCGCTTACCGGATACCTGTCGCGCTTTTCTCCCTTCGGGAG	3302

Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI CGAP Library."

```

BASE COUNT      230 a 205 c 214 g 182 t
ORIGIN
Query Match      14.8%; Score 780.8; DB 10; Length 831;
Best Local Similarity 99.3%; Pred. No. 3.2e-207;
Matches 826; Conservative 0; Mismatches 2; Indels 4; Gaps 4;
QY 536 GTTCTCTGTCGCCGCTGCTGCTCCGGTCTCTCGAGGGCCCGCGCGCGCG 595
DB 1 GTTCTCTGTCGCCGCTGCTGCTCCGGTCTCTCGAGGGCCCGCGCGCGCGCG 60
QY 596 CTTTCAGTCCGGACTCGACCTCTCGACGCGGAGCCGACGCGGGTGAGCCACGCTT 655
DB 61 CTTTCAGTCCGGACTCGACCTCTCGACGCGGAGCCGACGCGGGTGAGCCACGCTT 120
QY 656 ATGCAAGCAAGATCTGAGGAGCAGTTACGTTCTGTCTGTCCAGTGTAGTGAATCATGA 715
DB 121 ATGCAAGCAAGATCTGAGGAGCAGTTACGTTCTGTCTGTCCAGTGTAGTGAATCATGA 180
QY 716 CTGTACTCTACCGAATATTGGAATATTACAAAGTGTACAGTGTACAGTGAAGGAGGCTGC 775
DB 181 CTGTACTCTACCGAATATTGGAATATTACAAAGTGTACAGTGTACAGTGAAGGAGGCTGC 240
QY 776 AACATAACAGAGACAGCCCAACCTCAACTCAAGGACAGAGAGATATAAAATTTGCTG 835
DB 241 AACATAACAGAGACAGCCCAACCTCAACTCAAGGACAGAGAGATATAAAATTTGCTG 300
QY 836 CAGCAATTATAATACAGAGATCTTGAAGATATTGTAATAGTGGAGAAAGACTCAAT 895
DB 301 CAGCAATTATAATACAGAGATCTTGAAGATATTGTAATAGTGGAGAAAGACTCAAT 360
QY 896 GCATCCGACGAGGTGTATAGATGTGGGAGAGATTGGAGTCGACGACAAACCT 955
DB 361 GCATCCGACGAGGTGTATAGATGTGGGAGAGATTGGAGTCGACGACAAACCT 419
QY 956 TCTTTAAACCTCCATGTGTCTCCGTCTACAGATGTGGGGGTGTCTGCAATAGTGGGGGC 1015
DB 420 TCTTTAAACCTCCATGTGTCTCCGTCTACAGATGTGGGGGTGTCTGCAATAGTGGGGGC 479
QY 1016 TGCAGTGTATGACAGCAGCAGCTTACCTCAGCAGAGTTATTGAAATACAGTGC 1075
DB 480 TGCAGTGTATGACAGCAGCAGCTTACCTCAGCAGAGTTATTGAAATACAGTGC 539
QY 1076 CTCTCTCTCAAGGGCCCCAAACAGTAAACAATCAGTTTGTGCAATCACTTCTGCGGAT 1135
DB 540 CTCTCTCTCAAGGGCCCCAAACAGTAAACAATCAGTTTGTGCAATCACTTCTGCGGAT 599
QY 1136 GCATGCTTAAACTGGATGTTTACAGACAGTTTCAATTCATTTATAGAGTTCCTCCAG 1195
DB 600 GCATGCTTAAACTGGATGTTTACAGACAGTTTCAATTCATTTATAGAGTTCCTCCAG 659
QY 1196 CAACACTACACAGTGTGAGGACGACAAACAGACCTGCCCAACATTCATGTGGAATA 1255
DB 660 CAACACTACACAGTGTGAGGACGACAAACAGACCTGCCCAACATTCATGTGGAATA 718
QY 1256 ATCAGATCTGAGATGCTGCTGCTGAGAGATTTATGTTTCTCGATGCTGGAGTG 1315
DB 719 ATCAGATCTGAGATGCTGCTGCTGAGAGATTTATGTTTCTCGATGCTGGAGTG 777
QY 1316 ACTCAACAGATGGATTCCATGATCTGTGGACCAAAACAGGAGCTGGATGA 1367
DB 778 ACTCAACAGATGGATT-CATGACATCTGTGACCAAAACAGGAGCTGGATGA 828

```

RESULT 9
CA488579 LOCUS 858 bp mRNA linear EST 14-NOV-2002
DEFINITION AGENCOURT_10808789 MAPcL Homo sapiens cDNA clone IMAGE:6720351,
mRNA sequence.
ACCESSION CA488579

CA488579.1 GI:24950807
EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 858)
AUTHORS NIH-WGC <http://mgc.nci.nih.gov/>
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished

Contact: Robert Strausberg, Ph.D.
Email: egapbe@mail.nih.gov
Tissue Procurement: Kristi A. Eglund, Ira Pastan
CDNA Library Preparation: Invitrogen Corp
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLNL4279 row: 1 column: 15
High quality sequence stop: 674.

FEATURES
source

1..858
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6720351"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, HTERT-HNE1",
"LNCaP"
/lab_host="EMDH10B"
/clone_lib="MAPcL"
/notes="Vector: pCMV-SPORT6; Site 1: EcoRV; Site 2: Not I;
Subtracted with brain, liver, lung, kidney and muscle.
Directionally cloned. Priming method: oligo-dT. Average
insert size: 1800 bp. Library amplification: 26,000 fold.
Kristi A. Eglund, James J. Vincent, Robert Strausberg,
Bunker Lee and Ira Pastan: Discovery of new breast
cancer genes encoding membrane and secreted proteins.
Manuscript submitted."

BASE COUNT 257 a 205 c 199 g 195 t 2 others
ORIGIN

Query Match 14.8%; Score 780.8; DB 14; Length 858;
Best Local Similarity 99.0%; Pred. No. 3.3e-207;
Matches 816; Conservative 0; Mismatches 4; Indels 4; Gaps 3;
QY 883 AGAAGACTCAATGATGCTGAGTGTGTATAGATGTGGGAGAGTTGGAGTC 942
DB 1 AGAAGACTCAATGATGCTGAGTGTGTATAGATGTGGGAGAGTTGGAGTC 60
QY 943 GCAGCAACACCTCTTAAACCTCCATGTGTCTGCTACAGATGTGGGGTTGCTGC 1002
DB 61 GCAGCAACACCTCTTAAACCTCCATGTGTGTCTGCTACAGATGTGGGGTTGCTGC 120
QY 1003 AATAGTGGGGTGGAGTGCATGAACACAGCAGAGTACCTCAGCAAGACGTTATT 1062
DB 121 AATAGTGGGGTGGAGTGCATGAACACAGCAGAGTACCTCAGCAAGACGTTATT 180
QY 1063 GAATATACAGTGCCTCTCTCAAGCCCAACAGTAAACATCAGTTTCCCAATCAC 1122
DB 181 GAATATACAGTGCCTCTCTCAAGCCCAACAGTAAACATCAGTTTCCCAATCAC 240
QY 1123 ACTTCTCCGATGATGCTTAAACTGGATGTTTACAGCAAGTTCATTCCATTATGA 1182
DB 241 ACTTCTCCGATGATGCTTAAACTGGATGTTTACAGCAAGTTCATTCCATTATGA 300
QY 1183 CGTTCCCTGCCAGCAACACTACCAAGTGTGAGGAGCGAACAAGCTTCCCAATCAC 1242
DB 301 CGTTCCCTGCCAGCAACACTACCAAGTGTGAGGAGCGAACAAGCTTCCCAATCAC 360
QY 1243 TACATGTGGAATATCAGATCTGTCAGATGCTGCTCAGCAAGATTTTATGTTTCTCG 1302
DB 361 TACATGTGGAATATCAGATCTGTCAGATGCTGCTCAGCAAGATTTTATGTTTCTCG 420

QY 1303 GATCTCGAGTACTCAACAGATGATTCATGATCTGTGGACCAACAGAGCTG 1362
 Db 421 GATCTCGAGTACTCAACAGATGATTCATGATCTGTGGACCAACAGAGCTG 480
 QY 1363 GATCAAGAGACCTGTGAGTGTCTGAGAGCGGGCTTCGGCTTCGAGCTGTGGACCC 1422
 Db 481 GATCAAGAGACCTGTGAGTGTCTGAGAGCGGGCTTCGGCTTCGAGCTGTGGACCC 540
 QY 1423 CACAAAGAACTAGACAGAACTCATGCCAGTGTCTGTAAGAAACAAACTCTTCCCGAGC 1482
 Db 541 CACAAAGAACTAGACAGAACTCATGCCAGTGTCTGTAAGAAACAAACTCTTCCCGAGC 600
 QY 1483 CAATGTGGGGCCAAACGAGAAATTTGATGAACACATGCGAGTGTGTATGTAAGAAAC 1542
 Db 601 CAATGTGGGGCCAAACGAGAAATTTGATGAACACATGCGAGTGTGTATGTAAGAAAC 660
 QY 1543 TGCCCCAGAAATCAACCCCTTAAATCTCGGAAATGTGCTGTAATGTACAGAAAGTCCA 1602
 Db 661 TGCCCCAGAAATCAACCCCTTAAATCTCGGAAATGTGCTGTAATGTACAGAAAGTCCA 720
 QY 1603 CAGAAATGCTTGTAAAGAGAAAGAGTTCACCAACCAACATGCGAGTGTACAG-ACG 1661
 Db 721 CAGAAATGCTTGTAAAGAGAAAGAGTTCACCAACCAACATGCGAGTGTACAGAAAG 780
 QY 1662 GCCATGTACG-ACACCGCCAGAGGC--TTGTGAGCCAGATTTT 1702
 Db 781 GCCATGTACGAAACCGCCAGAAAGCCTTTGTGAGCCAGATTTT 824

RESULT 10
 CB686421/c
 LOCUS
 DEFINITION
 Bn01b_04j14.A 833 bp mRNA linear EST 09-APR-2003
 Bn01b_AAPC_EOORC transgenic Brassica napus overexpressing BNCBF17 c
 onstitutively frost tolerant Brassica napus cDNA clone Bn01b_04j14,
 mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Brassica napus (rape)
 Brassica napus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 1; eurosids II; Brassicales; Brassicaceae; Brassica.
 Singh, J., Allard, G., Tinker, N., Robert, L., Lacroix, C., De Moors, A.,
 Chagnon, J., Fatah, S., Couroux, P., and Hattori, J.
 Expressed Sequence Tags from constitutively frost tolerant
 transgenic Brassica napus overexpressing BNCBF17
 Unpublished
 Contact: Singh, J.A.
 Eastern Cereal and Oilseed Research Centre
 Agriculture and Agri-food Canada
 KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A
 OC6, Canada
 Tel: (613) 759-1662
 Fax: (613) 759-1701
 Email: singhja@agr.ca.
 Location/Qualifiers

1.
 /organism="Brassica napus"
 /mol_type="mRNA"
 /cultivar="Westar"
 /db_xref="taxon:3708"
 /clone="Bn01b_04j14"
 /tissue_type="fourth leaf"
 /dev_stage="3 weeks seedling grown at room temperature"
 /clone_lib="Bn01b_AAPC_EOORC transgenic Brassica napus_ove
 repressing BNCBF17 constitutively frost tolerant"
 /note="Vector: Bluescript SK+/XhoI-EcoRI; Site 1: EcoRI;
 Site 2: XhoI; Germinated in soil flats and seedlings grown
 for 3 weeks in a Conviron E-15 cabinet set at 20°C/16 hr

light (250 Em-2sec-1) and 16 °C / 8 hr dark. Fourth leaves
 collected at 9 am and immediately frozen." 1 others
 BASE COUNT 199 a 223 c 238 g 182 t
 ORIGIN
 Query Match
 Best Local Similarity 14.8%; Score 780.6; DB 14; Length 833;
 Matches 805; Conservative 99.0%; Pred. No. 3.6e-207;
 Mismatches 5; Indels 2; Gaps 2;
 QY 2701 GCTCTCTAGACGTAATCATGCTCATAGCTGTTTCCTGTGTGAATGTATCCGCTACA 2760
 Db 812 GCGCGCTTGGCGTAATCATGCTCATAGCTGTTTCCTGTGTGAATGTATCCGCTACA 753
 QY 2761 ATTCCACACAACTACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTTAATGAGTG 2820
 Db 752 ATTCCACACAACTACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTTAATGAGTG 693
 QY 2821 AGCTAACTCACTTAATTTGGTTTGGCTCAGCTCGCGCTTTCCAGTCCGGAAACCTGTG 2880
 Db 692 AGCTAACTCACTTAATTTGGTTTGGCTCAGCTCGCGCTTTCCAGTCCGGAAACCTGTG 633
 QY 2891 TGCAGCTGCAATTAATGAATCGGCCAAACCGCGGGGAGAGCGGTTTGCCTATTGGGGCGC 2940
 Db 632 TGCAGCTGCAATTAATGAATCGGCCAAACCGCGGGGAGAGCGGTTTGCCTATTGGGGCGC 573
 QY 2941 TCTTCCGCTTCCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCGCGAGCGGTA 3000
 Db 572 TCTTCCGCTTCCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCGCGAGCGGTA 513
 QY 3001 TCAGCTCACTCAAGGGGTATACGGTTATCCACAGATCAGGGGATTAACGCGAGGAAG 3060
 Db 512 TCAGCTCACTCAAGGGGTATACGGTTATCCACAGATCAGGGGATTAACGCGAGGAAG 453
 QY 3061 AACATGTGAGCAAAAGGCCAGAAAGGCCAGAAACCGTAAAGAGCGCGTTCCTGTCGCG 3120
 Db 452 AACATGTGAGCAAAAGGCCAGAAAGGCCAGAAACCGTAAAGAGCGCGTTCCTGTCGCG 393
 QY 3121 TTTTTCATAGGCTCCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCG 3180
 Db 392 TTTTTCATAGGCTCCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCG 333
 QY 3181 TGGGAAACCGGACGAGCTATAAGATACAGGGGTTTCCCTCTGGAAGCTCCCTGCTG 3240
 Db 332 TGGGAAACCGGACGAGCTATAAGATACAGGGGTTTCCCTCTGGAAGCTCCCTGCTG 273
 QY 3241 CGCTCTCTCTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGCGA 3300
 Db 272 CGCTCTCTCTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGCGA 213
 QY 3301 AGCGTGGCGCTTCTCATAGCTCAAGCTGATGATCTCAGTTGGTGTAGTGTGCTGCG 3360
 Db 212 AGCGTGGCGCTTCTCATAGCTCAAGCTGATGATCTCAGTTGGTGTAGTGTGCTGCG 153
 QY 3361 TCCAAAGTGGGCTGTGTGCAACACCCCGCTTCAGCCCGGAGCGCTGCGCTTAT-CCGG 3419
 Db 152 TCCAAAGTGGGCTGTGTGCAACACCCCGCTTCAGCCCGGAGCGCTGCGCTTAT-CCGG 93
 QY 3420 TAACATCTGCTTGTAGTCCAAACCGGTGAACACACCTATTCGACCTGGCAGAGCCAC 3479
 Db 92 TAACATCTGCTTGTAGTCCAAACCGGTGAACACACCTATTCGACCTGGCAGAG-CAC 34
 QY 3480 TGGTAAACAGATTAGCAGAGCGAGGTATGTAGG 3512
 Db 33 TGGTAAACAGATTAGCAGAGCGAGGTATGTAGG 1

RESULT 11
 AL044364
 LOCUS
 DEFINITION
 DXFP4343C172 s1 434 (synonym: htes3) Homo sapiens cDNA clone
 DXFP4343C172_3, mRNA sequence.
 ACCESSION
 AL044364
 VERSION
 AL044364.1 GI:5432586

EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 954)
 AUTHORS Ansoerge.W., Benes.V., Krieger.S., Mewes.H.W., Gassenhuber,J. and
 Wiemann.S.
 TITLE EST (Ansoerge, Benes, et al.)
 JOURNAL Unpublished
 COMMENT Contact: Ansoerge W
 MIPS
 Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
 This is the 3' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by EMBL (European Molecular Biology Laboratories,
 Heidelberg/Germany) within the CDNA sequencing consortium of the
 German Genome Project.
 No r1 sequence available.
 This clone (DKFZp343C172) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
 FEATURES
 source
 1..954
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp343C172"
 /tissue_type="testis"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="434 (synonym: htes3)"
 /note="vector: pSPori1; Site_1: NotI; Site_2: SalI"
 BASE COUNT 214 a 273 c 253 g 214 t
 ORIGIN
 Query Match 14.7%; Score 774; DB 9; Length 954;
 Best Local Similarity 100.0%; Pred. No. 2.8e-205;
 Matches 774; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 2887 CTGCATTAAATGAAATCGGCCAACCGCGGGGAGAGCGGTTGGGTATGGCGCTTTTC 2946
 Db 181 CTGCATTAAATGAAATCGGCCAACCGCGGGGAGAGCGGTTGGGTATGGCGCTTTTC 240
 2947 GCTTCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 3006
 Db 241 GCTTCCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCA 300
 3007 CACTCAAGGCGGTAATACGTTTATCCACAGAAATCAGGGGATACCGAGAAAGACATG 3066
 Db 301 CACTCAAGGCGGTAATACGTTTATCCACAGAAATCAGGGGATACCGAGAAAGACATG 360
 3067 TGAGCAAAAGCCAGCAAAAGCCAGGACCGTAAAGCCGCGGTTGCTGGCGTTTTC 3126
 Db 361 TGAGCAAAAGCCAGCAAAAGCCAGGACCGTAAAGCCGCGGTTGCTGGCGTTTTC 420
 3127 CATAGGCTCCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGTGGCGA 3186
 Db 421 CATAGGCTCCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGTGGCGA 480
 3187 AACCCGACGAGCTATAAGATACGAGCGTTTCCCTGGAAGCTCCCTCGTGGCTCT 3246
 Db 481 AACCCGACGAGCTATAAGATACGAGCGTTTCCCTGGAAGCTCCCTCGTGGCTCT 540
 3247 CTGTTTCGACCCCTCGCCCTTACCGGATACCTGTGCGCTTTCTCCCTTCGGGAAGCGTG 3306
 Db 541 CTGTTTCGACCCCTCGCCCTTACCGGATACCTGTGCGCTTTCTCCCTTCGGGAAGCGTG 600
 3307 GCGCTTTTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAG 3366
 Db 601 GCGCTTTTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAG 660

3367 CTGGGCTGTGTGCACGAACCCCGGTTAGCCGAGCGCTTGGCGCTTATCCGGTAACTAT 3426
 Db 661 CTGGGCTGTGTGCACGAACCCCGGTTAGCCGAGCGCTTGGCGCTTATCCGGTAACTAT 720
 3427 CGTCTTGGTCCAAACCCGGTAGACAGCACTTATGCCACTGGCAGCAGCACTGGTAAC 3486
 Db 721 CGTCTTGGTCCAAACCCGGTAGACAGCACTTATGCCACTGGCAGCAGCACTGGTAAC 780
 3487 AGGATTAGCAGACGAGGATATGAGCGGTGTACAGAGTTCTTGAAGTGTGGCTTAAC 3546
 Db 781 AGGATTAGCAGACGAGGATATGAGCGGTGTACAGAGTTCTTGAAGTGTGGCTTAAC 840
 3547 TAGGGCTACTAGAAACAGATTTTGGTATCTGCGCTCTGCTGAAGCAGTTACTTTC 3606
 Db 841 TAGGGCTACTAGAAACAGATTTTGGTATCTGCGCTCTGCTGAAGCAGTTACTTTC 900
 3607 GGAATAAGATTGGTAGCTTTGATCCGCAAAACAAACCCCTGGTAGCGT 3660
 Db 901 GGAATAAGATTGGTAGCTTTGATCCGCAAAACAAACCCCTGGTAGCGT 954
 RESULT 12
 BZ579291 1369 bp DNA linear GSS 17-DEC-2002
 LOCUS msh2_6285.x1 msh Pseudomonas aeruginosa genomic clone msh2_6285,
 DEFINITION genomic survey sequence.
 ACCESSION BZ579291
 VERSION BZ579291.1 GI:27214352
 KEYWORDS GSS.
 SOURCE Pseudomonas aeruginosa
 ORGANISM Pseudomonas aeruginosa
 Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 Pseudomonadaceae; Pseudomonas.
 REFERENCE 1 (bases 1 to 1369)
 AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
 Burns,J.L., Kaul,R. and Olsen,M.V.
 TITLE Whole-Genome-Sequence variation among multiple isolates of
 Pseudomonas aeruginosa library
 JOURNAL J. Bacteriol., (2002) In press
 COMMENT Contact: Chris K. Raymond
 Genome Center
 University of Washington
 Box 352145, Seattle, WA 98105-2145, USA
 Tel: 2062216954
 Fax: 2066857244
 Email: craymond@u.washington.edu
 Class: shotgun.
 Location/Qualifiers
 1..1369
 /organism="Pseudomonas aeruginosa"
 /mol_type="genomic DNA"
 /strains="MSH"
 /db_xref="taxon:287"
 /clone="msh2_6285"
 /clone_lib="msh"
 /note="Environmental isolate. Whole genomic shotgun
 library."
 BASE COUNT 327 a 356 c 351 g 334 t 1 others
 ORIGIN
 Query Match 14.3%; Score 754; DB 29; Length 1369;
 Best Local Similarity 95.6%; Pred. NO.1.5e-199;
 Matches 797; Conservative 0; Mismatches 35; Indels 2; Gaps 2;
 2840 CGTTGGCTCACTGCCCGCTTCCAGTCGGAAACCTTCGTCGAGTGATTAATGAA 2899
 Db 99 CGTTGGCTCACTGCCCGCTTCCAGTCGGAAACCTTCGTCGAGTGATTAATGAA 158
 2900 TCGGGCCAAACGCGGGGAGAGCGGTTTGGTATTGGCGCTCTTTCCGCTTCCTCGCTCA 2959
 Db 159 TCGGGCCAAACGCGGGGAGAGCGGTTTGGTATTGGCGCTCTTTCCGCTTCCTCGCTCA 218
 2960 CTGACTCGTTCGCTCGGTCGCTTCGGCTTCGGCGAGCGGTATCAGTCACTCAAGGCGG 3019

LOCUS B2572730 1346 bp DNA linear GSS 17-DEC-2002
 DEFINITION msh2_2771.x1 msh Pseudomonas aeruginosa genomic clone msh2_2771,
 genomic survey sequence.
 ACCESSION B2572730
 VERSION B2572730.1 GI:27207791
 KEYWORDS GSS.
 SOURCE Pseudomonas aeruginosa
 ORGANISM Pseudomonas aeruginosa
 Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 Pseudomonadaceae; Pseudomonas.
 REFERENCE 1 (bases 1 to 1346)
 AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
 Burns,J.L., Kaul,R. and Olsen,M.V.
 TITLE Whole-Genome-Sequence variation among multiple isolates of
 Pseudomonas aeruginosa library
 JOURNAL Pseudomonas aeruginosa library
 COMMENT J. Bacteriol., (2002) in press
 Contact: Chris K. Raymond
 Genome Center
 University of Washington
 Box 352145, Seattle, WA 98105-2145, USA
 Tel: 2062216954
 Fax: 2066857244
 Email: craymond@u.washington.edu
 Class: shotgun.
 FEATURES
 source
 1. 1346
 Location/Qualifiers
 /organism="Pseudomonas aeruginosa"
 /mol_type="genomic DNA"
 /strain="MSH"
 /db_xref="taxon:287"
 /clone_lib="msh2_2771"
 /notes="msh"
 /notes="Environmental isolate. Whole genomic shotgun
 library."
 BASE COUNT 299 a 369 c 346 g 330 t 2 others
 ORIGIN
 Query Match 14.0%; Score 742.2; DB 29; Length 1346;
 Best Local Similarity 87.9%; Pred. No. 3e-196;
 Matches 855; Conservative 0; Mismatches 110; Indels 8; Gaps 4;
 QY 2835 AATTGCGTTGCGCTCACTGCGCGCTTCCAGTCGGGAACCTGTCGTCGAGTGCATTA 2894
 Db 84 AATTCTGCGAGCCCACTGCCTGTTTCCAGTCGGGAACCTGTCGTCGAGTGCATTA 143
 QY 2895 ATGAATCGGCACCGCGGGGAGAGCGGTTTCGATTATTTGGGCGCTCTTCGCGTTCCCTC 2954
 Db 144 ATGAATCGGCACCGCGGGGAGAGCGGTTTCGATTATTTGGGCGCTCTTCGCGTTCCCTC 203
 QY 2955 GCTCACTGACTCGCTGCGCTCGGTCGTTTCGCTCGCGCGAGCGGTATCAGCTCACTCAA 3014
 Db 204 GCTCACTGACTCGCTGCGCTCGGTCGTTTCGCTCGCGCGAGCGGTATCAGCTCACTCAA 263
 QY 3015 GCGGTTAATACGGTTATCCAGAAATCAGGGGATAACGCGAGGAAGAACATGTGAGCAA 3074
 Db 264 GCGGTTAATACGGTTATCCAGAAATCAGGGGATAACGCGAGGAAGAACATGTGAGCAA 323
 QY 3075 AGGCCACAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGCGCTTTTCCATAGGCT 3134
 Db 324 AGGCCACAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGCGCTTTTCCATAGGCT 383
 QY 3135 CCGCCCCCTTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGAC 3194
 Db 384 CCGCCCCCTTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGAC 443
 QY 3195 AGGACTATAAGATACGAGCGTTTCCCTCGGAAGCTCCCTCGTCGCTCTCTCTGTTC 3254
 Db 444 AGGACTATAAGATACGAGCGTTTCCCTCGGAAGCTCCCTCGTCGCTCTCTCTGTTC 503
 QY 3255 GACCTGCGGTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGAAGCGTGGCGCTTC 3314
 Db 504 GACCTGCGGTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGAAGCGTGGCGCTTC 563

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 19:10:26 ; Search time 1979.02 Seconds
(without alignments)
10335.846 Million cell updates/sec

Title: US-09-921-143-36_COPY_1_500
Perfect score: 500
Sequence: 1 aagctgacattatgcact.....cactgggtggatgcacg 500

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 2045481386 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_ov.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	182	36.4	1560	14	AELENV K02374 Rous-associ
2	175	35.0	9396	14	AF052428 AF052428 Rous eac
3	140	28.0	346	14	AELETR2 K00993 Rous-associ
4	127	25.4	281	6	BD136066 BD136066 Novel pro
5	127	25.4	282	6	BD136065 BD136065 Novel pro
6	127	25.4	512	6	BD136064 BD136064 Novel pro
7	127	25.4	599	6	E59416 E59416 Signal pept
8	127	25.4	599	6	E51990 E51990 Promoter, r
9	127	25.4	600	6	AX360741 AX360741 Sequence
10	127	25.4	619	6	AX256412 AX256412 Sequence
11	127	25.4	738	6	AX591955 AX591955 Sequence
12	127	25.4	777	6	BD000145 BD000145 Vector ha
13	127	25.4	777	6	I05430 I05430 Sequence 16
14	127	25.4	777	6	I08105 I08105 Sequence 3
15	127	25.4	875	14	HS5PIE1AB M64941 Human cytom
16	127	25.4	876	14	HS5PIE1AE M64944 Human cytom
17	127	25.4	876	14	HS5PIE1AF M64940 Human cytom
18	127	25.4	877	14	HS5PIE1AC M64942 Human cytom
19	127	25.4	919	6	BD000143 BD000143 Vector ha
20	127	25.4	919	6	I05393 I05393 Sequence 14
21	127	25.4	919	6	I08103 I08103 Sequence 1
22	127	25.4	930	6	A10617 A10617 Human Cytom
23	127	25.4	930	6	AR028792 AR028792 Sequence
24	127	25.4	930	6	AR028794 AR028794 Sequence
25	127	25.4	930	6	AR050544 AR050544 Sequence
26	127	25.4	930	6	AR050546 AR050546 Sequence
27	127	25.4	930	6	AR094363 AR094363 Sequence
28	127	25.4	930	6	AR241121 AR241121 Sequence
29	127	25.4	930	6	AR241123 AR241123 Sequence
30	127	25.4	930	6	BD131785 BD131785 Transfect
31	127	25.4	930	6	BD145110 BD145110 Methods a
32	127	25.4	930	6	BD145112 BD145112 Methods a
33	127	25.4	930	6	E00836 E00836 Recombinant
34	127	25.4	930	6	I49834 I49834 Sequence 2
35	127	25.4	930	6	I49836 I49836 Sequence 4
36	127	25.4	930	14	HS5IEE K03104 Human cytom
37	127	25.4	996	6	AX282811 AX282811 Sequence
38	127	25.4	1070	6	A85308 A85308 Sequence 6
39	127	25.4	1070	6	BD107647 BD107647 FIV vacci
40	127	25.4	1078	6	A92081 A92081 Sequence 5
41	127	25.4	1078	6	AR230688 AR230688 Sequence
42	127	25.4	1318	6	A92080 A92080 Sequence 4
43	127	25.4	1318	6	AR230690 AR230690 Sequence
44	127	25.4	1417	6	A92077 A92077 Sequence 1
45	127	25.4	1417	6	AR230692 AR230692 Sequence

ALIGNMENTS

RESULT 1

AELENV

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

AELENV

Rous-associated virus-2 env and gag genes.

K02374.1

env protein; gag protein; long terminal repeat (LTR).

Rous-associated virus type 2

Rous-associated virus type 2

Viruses; Retrovirus; Retroviridae; Alpharetrovirus.

1 (bases 1 to 1560)

Bizub,D., Katz,R.A. and Skalka,A.M.

Nucleotide sequence of noncoding regions in Rous-associated

virus-2: comparisons delineate conserved regions important in

replication and oncogenesis

1560 bp ss-RNA linear

VRL 28-APR-1993


```

mat_peptide      2130..2501
/gene="gag"
/product="p15"
<2522..5210
/gene="pol"
<2522..5209
/gene="pol"
/codon_start=1
/product="polymerase and integrase protein"
/protein_id="AAC08988.1"
/db_xref="GI:3003002"
/translation="TVALLHAIPLKWKPHPTPWIDOWLPPEGLKVALTOLVEKELQ
LHTEPSLQNTWTFVIRKASGYSRLHDLRAVNAKLPFGAVQGGAPVLSALPGRWP
LVMLDKDFPFPLAEQDREAFPLPSPNQAPARRQFWKVLPGMTCPTICQLV
VGVLEPLRLKPSLRLMVDLLAAASHDGLAAGEVINTLERAGFTTSPKIQ
REGVOYLGRKLGSTVAVPVGLVAEPRIATLDVOKLVSLOWLPALGIPRLMGPF
YEQLRSDPNEAREWNLDMKMAWEIYVLSLTAALRMDPAPLGLGAVRCQGAIV
LGGDLPLPFCILALRGFAKINSSTPFIIDARPLVSLKRVTDHPVPTFTD
FREDLPLPFCILALRGFAKINSSTPFIIDARPLVSLKRVTDHPVPTFTD
ASSTHKGVVVRGPRWEIKETADIGASVQOLEARAVAMALLMPTFTTNVVTDSAF
VAKMLLMKQGGVPTAAAFILDALQSASMAVHLVRSHEVFCFTEGNDVADSQ
ATFOALVREAKLHTTLHIGPRALSKACNISMQOAREVVQTCPCNSAPALEAGVNP
RGLGPIQWQDFTLPRMAPRSMLAVTVDASSAIVTQHGRTVSVAQGHWAIA
VLGRPAIKTDNGSCSTSKSTREWLARWCIATHTGIPQNSQOAMVERANRLKIKR
VLAEGDGMKRIIPASQOGLAKAMTALNHFERGENTKTPVQKHWRPVLITGPPVPI
RIETGWEKGNVLMVWGRYAAVKNRDTDKVLPVSRKVPDITQKDEVTKKDRASPL
FAGSSDWIPMGDEQGLQEEAASNKQEGFCEDTLAANES"
6961..7040
/rpt_type=direct
7300..8908
/gene="src"
7300..8907
/gene="src"
/codon_start=1
/product="src tyrosine kinase"
/protein_id="AAC08990.1"
/db_xref="GI:3003004"
/translation="MSSSKSKPKDPHSRRSLRLEPPDSTHGGFPASQTPNKTAAPDTH
RTPSRFGTAVTEKFLGDFNTSDTVSPQAGALAGGVTTFVALDYDSEWLETDLSP
KGERLQIVNNTGNNWLAHSVTGTGIVSPNYAPSDSIQAEWYFGKIRRESER
LLANPRTGTFVLRSETHGTGTSVDFDNAGLVNHYKIRKLDGSGFYITSR
QFSSLOLVAYYSKHADGLCHRLTNVCPTSPQOTGLAKDAWEIPRESLRLEKVLQ
CFGEVWNGTGTGTGTGTGTVAIKTLKPGTMSPEAPLOBAOVAKLREKVLQ
YAVVSEPIVILVEYMSKSLDLFLKEMGKYLRLPOLDMAAQIATSGMAVVERNNYV
HRLRANILVGENLVCKVADFLGALLEDNEYTARQCAKFIKWTAPAAIYGRFTI
KSDVWSFGILLTETLTKGRFPMGMNGEVLDRVERGYRMPCPPECPSLHDLMCQQW
RRDPERPFTFXYLQALLPACVLEVAE"
9002..9096
/rpt_type=direct
9131..9396
9131..9375
/note="U3"
9369..9374
9376..9396
/note="R"

BASE COUNT      2226 a 2366 c 2740 g 2064 t
ORIGIN
Query Match      35.0%; Score 175; DB 14; Length 9396;
Best Local Similarity 100.0%; Pred. No. 5e-95;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 259 TGAGTTAGCAACATGCCCTTATAGGAGAGAAAAAGCACCGTGCATGCCGATTGGTGGGAG 318
Db 9183 TGAGTTAGCAACATGCCCTTATAGGAGAGAAAAAGCACCGTGCATGCCGATTGGTGGGAG 9242

Qy 319 TAAGTGGTATGATCGTGGTATGATCGTGCCTTGTAGGAAGGCAACAGACGGGTCTAAC 378
Db 9243 TAAGTGGTATGATCGTGGTATGATCGTGCCTTGTAGGAAGGCAACAGACGGGTCTAAC 9302

Qy 379 ACGGATTGGAGCAACCACTGAATTCGCGATTGCAGAGATATTCGTTAAGTGCC 433
Db 9303 ACGGATTGGAGCAACCACTGAATTCGCGATTGCAGAGATATTCGTTAAGTGCC 9357

```

```

RESULT 3
ALELTR2
LOCUS
DEFINITION
endonuclease cleavage sites.
ACCESSION
K00993
VERSION
GI:210075
KEYWORDS
LTR; reverse transcriptase endonuclease cleavage site.
SOURCE
Rous-associated virus type 2
ORGANISM
Rous-associated virus type 2
VIRUSES; Retrovirdae; Alpharetrovirus.
REFERENCE
1 (bases 1 to 346)
Duyk, G., Leis, J., Longiaru, M. and Skalka, A.M.
Selective cleavage in the avian retroviral long terminal repeat
sequence by the endonuclease associated with the alpha beta form of
avian reverse transcriptase
Proc. Natl. Acad. Sci. U.S.A. 80 (22), 6745-6749 (1983)
84070706
MEDLINE
6196775
PUBMED
2 (bases 1 to 346)
Skalka, A.M., Duyk, G., Longiaru, M., DeHaseth, P., Terry, R. and
Leis, J.
Integrative recombination--a role for the retroviral reverse
transcriptase
Cold Spring Harb. Symp. Quant. Biol. 49, 651-659 (1984)
85153042
MEDLINE
6085046
PUBMED
COMMENT
Original source text: Rous-associated virus 2 replication form I
DNA, clone RAV2-2 [1], and plasmids pP01 and pGJ14 [2].
Reverse transcriptase associated endonuclease (purified from avian
sarcoma virus) cleavage sites have been mapped in two tandemly
linked Rous-associated virus-2 LTR sequences. The enzyme may be
involved in viral cDNA integration in the host, since it generates
a 6 bp staggered overlap that spans the junction. The clone
sequence (RAV2-2) corresponds to the unintegrated replicative form
(RF) I of RAV-2.
Draft entry for [2] kindly provided by A. Skalka, 15-AUG-1985.
Location/Qualifiers
1..346
/organism="Rous-associated virus type 2"
/mol_type="genomic RNA"
/db_xref="taxon:11948"
<1..154
/note="LTR A"
complement(104..105)
/note="rt-endonuclease cleavage site"
130..131
/note="rt-endonuclease secondary cleavage site"
151..152
/note="rt-endonuclease primary cleavage site"
155..2346
/note="LTR B"
complement(157..158)
/note="rt-endonuclease primary cleavage site"
214..215
/note="rt-endonuclease secondary cleavage site"
complement(222..223)
/note="rt-endonuclease secondary cleavage site"
complement(244..245)
/note="rt-endonuclease secondary cleavage site"
304..305
/note="rt-endonuclease secondary cleavage site"
complement(306..307)
/note="rt-endonuclease secondary cleavage site"
BASE COUNT      96 a 69 c 89 g 92 t
ORIGIN
EcORI site.
Query Match      28.0%; Score 140; DB 14; Length 346;
Best Local Similarity 100.0%; Pred. No. 1.5e-73;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 259 TGAGTTAGCAACATGCCCTTATAGGAGAGAAAAAGCACCGTGCATGCCGATTGGTGGGAG 318

```

```

Db      207 TGAGTTAGCAATCCCTTATAAGGAGAGAAAACACCGTGCATGCGGATGCTGGGAG 266
QY      319 TAAGTGGTATGATCGTGTATGATCGCTTGTAGGAGGCAACAGACGGGTCTAAC 378
Db      267 TAAGTGGTATGATCGTGTATGATCGCTTGTAGGAGGCAACAGACGGGTCTAAC 326
QY      379 ACGGATTGACGAACCACTG 398
Db      327 ACGGATTGACGAACCACTG 346

RESULT 4
LOCUS   BD136066
DEFINITION Novel promoter element for sustained gene expression.
ACCESSION BD136066
VERSION   BD136066.1 GI:23231011
KEYWORDS JP 2002508974-A/3.
SOURCE   Cytomegalovirus
ORGANISM Cytomegalovirus
VIRUSES; dsDNA viruses, no RNA stage; Herpesviridae;
Betaherpesvirinae.
REFERENCE 1 (bases 1 to 281)
AUTHORS Armentano,D., Yew,N. and Marshall,J.
TITLE Novel promoter element for sustained gene expression
JOURNAL Patent: JP 2002508974-A 3 26-MAR-2002;
GENZYME CORP
COMMENT OS Cytomegalovirus
PN JP 2002508974-A/3
PD 26-MAR-2002
PF 15-JAN-1999 JP 2000540258
PR 16-JAN-1998 US 60/071673
PI DONNA ARMENTANO,NELSON YEW,JOHN MARSHALL
PC C12N15/09,A61K35/76,A61K48/00,C12N15/00
CC Novel promoter element for sustained gene expression FH Key
Location/Qualifiers
FT source 1..281
FT /organism='Cytomegalovirus'.

Query Match 25.4%; Score 127; DB 6; Length 281;
Best Local Similarity 100.0%; Pred. No. 1.4e-65;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACATGGTGATCGG 76
Db 31 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACATGGTGATCGG 90
QY 77 TTTTGGCAGTACATCAATGGCGGTGATAGCGTTTGACTCACGGGATTTCCAAAGTCTC 136
Db 91 TTTTGGCAGTACATCAATGGCGGTGATAGCGTTTGACTCACGGGATTTCCAAAGTCTC 150
QY 137 CACCCCA 143
Db 151 CACCCCA 157

RESULT 5
LOCUS   BD136065
DEFINITION Novel promoter element for sustained gene expression.
ACCESSION BD136065
VERSION   BD136065.1 GI:23231010
KEYWORDS JP 2002508974-A/2.
SOURCE   Cytomegalovirus
ORGANISM Cytomegalovirus
VIRUSES; dsDNA viruses, no RNA stage; Herpesviridae;
Betaherpesvirinae.
REFERENCE 1 (bases 1 to 281)
AUTHORS Armentano,D., Yew,N. and Marshall,J.
TITLE Novel promoter element for sustained gene expression
JOURNAL Patent: JP 2002508974-A 2 26-MAR-2002;
GENZYME CORP
COMMENT OS Cytomegalovirus
PN JP 2002508974-A/2
PD 26-MAR-2002
PF 15-JAN-1999 JP 2000540258
PR 16-JAN-1998 US 60/071673
PI DONNA ARMENTANO,NELSON YEW,JOHN MARSHALL
PC C12N15/09,A61K35/76,A61K48/00,C12N15/00
CC Novel promoter element for sustained gene expression FH Key
Location/Qualifiers
FT source 1..282
FT /organism='Cytomegalovirus'.

Query Match 25.4%; Score 127; DB 6; Length 282;
Best Local Similarity 100.0%; Pred. No. 1.4e-65;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACATGGTGATCGG 76
Db 27 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACATGGTGATCGG 86
QY 77 TTTTGGCAGTACATCAATGGCGGTGATAGCGTTTGACTCACGGGATTTCCAAAGTCTC 136
Db 87 TTTTGGCAGTACATCAATGGCGGTGATAGCGTTTGACTCACGGGATTTCCAAAGTCTC 146
QY 137 CACCCCA 143
Db 147 CACCCCA 153

RESULT 6
LOCUS   BD136064
DEFINITION Novel promoter element for sustained gene expression.
ACCESSION BD136064
VERSION   BD136064.1 GI:23231009
KEYWORDS JP 2002508974-A/1.
SOURCE   Cytomegalovirus
ORGANISM Cytomegalovirus
VIRUSES; dsDNA viruses, no RNA stage; Herpesviridae;
Betaherpesvirinae.
REFERENCE 1 (bases 1 to 512)
AUTHORS Armentano,D., Yew,N. and Marshall,J.
TITLE Novel promoter element for sustained gene expression
JOURNAL Patent: JP 2002508974-A 1 26-MAR-2002;
GENZYME CORP
COMMENT OS Cytomegalovirus
PN JP 2002508974-A/1
PD 26-MAR-2002
PF 15-JAN-1999 JP 2000540258
PR 16-JAN-1998 US 60/071673
PI DONNA ARMENTANO,NELSON YEW,JOHN MARSHALL
PC C12N15/09,A61K35/76,A61K48/00,C12N15/00
CC Novel promoter element for sustained gene expression FH Key
Location/Qualifiers
FT source 1..512
FT /organism='Cytomegalovirus'.

```

```

/mol_type="genomic DNA"
/db_xref="taxon:10358"
BASE COUNT      126 a 131 c 123 g 132 t
ORIGIN
Query Match      25.4%; Score 127; DB 6; Length 512;
Best Local Similarity 100.0%; Pred. No. 1.4e-65;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 17 GAC TTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGATGCGG 76
Db 257 GAC TTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGATGCGG 316
Qy 77 TTTTGGCAGTACATCAATGGCGGTGATACCGGTTTGACTCAGCGGGATTTCCAAGTCTC 136
Db 317 TTTTGGCAGTACATCAATGGCGGTGATACCGGTTTGACTCAGCGGGATTTCCAAGTCTC 376
Qy 137 CACCCCA 143
Db 377 CACCCCA 383

RESULT 7
E59416
LOCUS      E59416
DEFINITION Signal peptide.
ACCESSION E59416
VERSION    E59416.1 GI:18622549
KEYWORDS   JP 2000354490-A/3.
SOURCE     unidentified
ORGANISM   unclassified.
REFERENCE  1 (bases 1 to 590)
AUTHORS    Hawkins,R. and Nakamura,M.
TITLE      Signal peptide
JOURNAL    Patent: JP 2000354490-A 3 26-DEC-2000;
COMMENT    OS cauliflower mosaic virus promoter
          PN JP 2000354490-A/3
          PD 26-DEC-2000
          PF 15-JUN-1999 JP 1999168271
          PR RICHARD HAWKINS,MICHI NAKAMURA
          PI C12N15/09,C07K14/61,C12N1/13,C12P21/02,C12N15/00 CC
          PC Key Location/Qualifiers
          FH source 1..590
          FT source /organism="cauliflower mosaic virus promoter".
FEATURES
  source
  Location/Qualifiers
    1..590
    /organism="unidentified"
    /mol_type="genomic DNA"
    /db_xref="taxon:32644"
BASE COUNT      149 a 143 c 141 g 157 t
ORIGIN

Query Match      25.4%; Score 127; DB 6; Length 590;
Best Local Similarity 100.0%; Pred. No. 1.4e-65;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 17 GAC TTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGATGCGG 76
Db 314 GAC TTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGATGCGG 373
Qy 77 TTTTGGCAGTACATCAATGGCGGTGATACCGGTTTGACTCAGCGGGATTTCCAAGTCTC 136
Db 374 TTTTGGCAGTACATCAATGGCGGTGATACCGGTTTGACTCAGCGGGATTTCCAAGTCTC 433
Qy 137 CACCCCA 143
Db 434 CACCCCA 440

RESULT 8
E51990
LOCUS      E51990
DEFINITION Promoter, recombinant containing the same and utilization thereof.
ACCESSION E51990
VERSION    E51990.1 GI:18629551
KEYWORDS   JP 2001000188-A/5.
SOURCE     Human herpesvirus 5
ORGANISM   Human herpesvirus 5
REFERENCE  1 (bases 1 to 599)
AUTHORS    Kubomura,M. and Saito,S.
TITLE      Promoter, recombinant containing the same and utilization thereof
JOURNAL    Patent: JP 2001000188-A 5 09-JAN-2001;
COMMENT    NIPPON ZEON CORP
          OS Cytomegalovirus
          PN JP 2001000188-A/5
          PD 09-JAN-2001
          PF 22-JUN-1999 JP 1999174804
          PR MAYUMI KUBOMURA,SHUJI SAITO
          PI C12N15/09,A61K31/00,A61K31/00,A61K39/02,A61K39/17,A61K39/215,
          PC C07K14/125,
          CC C07K14/165,C07K14/30,C12N7/00,C12N15/00
          FH Key Location/Qualifiers
          FT source 1..599
          FT /organism="Cytomegalovirus".
FEATURES
  source
  Location/Qualifiers
    1..599
    /organism="Human herpesvirus 5"
    /mol_type="genomic DNA"
    /db_xref="taxon:10359"
BASE COUNT      152 a 144 c 144 g 159 t
ORIGIN

Query Match      25.4%; Score 127; DB 6; Length 599;
Best Local Similarity 100.0%; Pred. No. 1.4e-65;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 17 GAC TTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGATGCGG 76
Db 319 GAC TTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGATGCGG 378
Qy 77 TTTTGGCAGTACATCAATGGCGGTGATAGCGGTTTGACTCAGCGGGATTTCCAAGTCTC 136
Db 379 TTTTGGCAGTACATCAATGGCGGTGATAGCGGTTTGACTCAGCGGGATTTCCAAGTCTC 438
Qy 137 CACCCCA 143
Db 439 CACCCCA 445

RESULT 9
AX360741
LOCUS      AX360741
DEFINITION Sequence 425 from Patent WO0202623.
ACCESSION AX360741
VERSION    AX360741.1 GI:18676241
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  1
AUTHORS    Wang,T., McNeill,P.D., Wantanabe,Y., Carter,D., Henderson,R.A. and
          Kalos,M.D.
TITLE      Compositions and methods for the therapy and diagnosis of lung
          cancer
JOURNAL    Patent: WO 0202623-A 425 10-JAN-2002;
          CORIXA CORPORATION (US)
FEATURES
  source
  Location/Qualifiers
    1..600

```



```

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 153 a 152 c 141 g 154 t
ORIGIN
Query Match 25.4%; Score 127; DB 6; Length 600;
Best Local Similarity 100.0%; Pred. No. 1.4e-65;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 GACATTTCTTCTTGGCAGTACATCTACGTATTAGTCTGCTATTACCATGGTGTATCGGG 76
Db 278 GACATTTCTTCTTGGCAGTACATCTACGTATTAGTCTGCTATTACCATGGTGTATCGGG 337
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCGAAGTCTC 136
Db 338 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCGAAGTCTC 397
QY 137 CACCCCA 143
Db 398 CACCCCA 404
RESULT 10
AX256412 619 bp DNA linear PAT 10-OCT-2001
LOCUS
DEFINITION Sequence 60 from Patent WO0170816.
ACCESSION AX256412
VERSION AX256412.1 GI:16075237
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE
AUTHORS Palli,S.R., Kapitskaya,M.Z. and Cress,D.E.
TITLE Ecdysone receptor-based inducible gene expression system
JOURNAL Patent: WO 0170816-A 60 27-SEP-2001;
ROHM AND HAAS COMPANY (US)
FEATURES
source
1..619
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
BASE COUNT 150 a 164 c 155 g 150 t
ORIGIN
Query Match 25.4%; Score 127; DB 6; Length 619;
Best Local Similarity 100.0%; Pred. No. 1.4e-65;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 GACATTTCTTCTTGGCAGTACATCTACGTATTAGTCTGCTATTACCATGGTGTATCGGG 76
Db 254 GACATTTCTTCTTGGCAGTACATCTACGTATTAGTCTGCTATTACCATGGTGTATCGGG 313
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCGAAGTCTC 136
Db 314 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCGAAGTCTC 373
QY 137 CACCCCA 143
Db 374 CACCCCA 380
RESULT 11
AX591955 738 bp DNA linear PAT 27-JAN-2003
LOCUS
DEFINITION Sequence 24 from Patent WO0236760.
ACCESSION AX591955
VERSION AX591955.1 GI:27950185
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Lin,J., Yaver,D., Foster,D. and Holly,R.
AUTHORS Methods for expressing endogenous genes by restriction enzyme
TITLE mediated integration
JOURNAL Patent: WO 0236760-A 24 10-MAY-2002;
Novozymes Biotech, Inc. (US); ZymoGenetics, Inc. (US)
FEATURES
Location/Qualifiers
1..738
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
BASE COUNT 188 a 178 c 168 g 204 t
ORIGIN
Query Match 25.4%; Score 127; DB 6; Length 738;
Best Local Similarity 100.0%; Pred. No. 1.4e-65;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 GACATTTCTTCTTGGCAGTACATCTACGTATTAGTCTGCTATTACCATGGTGTATCGGG 76
Db 335 GACATTTCTTCTTGGCAGTACATCTACGTATTAGTCTGCTATTACCATGGTGTATCGGG 394
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCGAAGTCTC 136
Db 395 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCGAAGTCTC 454
QY 137 CACCCCA 143
Db 455 CACCCCA 461
RESULT 12
BD000145 777 bp DNA linear PAT 31-JAN-2002
LOCUS
DEFINITION Vector having stabilized sequence and eucaryotic host cell.
ACCESSION BD000145
VERSION BD000145.1 GI:18623224
KEYWORDS JP 2000308497-A/3.
SOURCE unidentified
ORGANISM unidentified
REFERENCE
1 (bases 1 to 777)
AUTHORS Goman,C.M.
TITLE Vector having stabilized sequence and eucaryotic host cell
JOURNAL Patent: JP 2000308497-A 3 07-NOV-2000;
GENETIC INC
COMMENT
CS Unidentified
FN JP 2000308497-A/3
PD 07-NOV-2000
PF 17-APR-2000 JP 2000115248
PR 12-SEP-1986 US 907185.09-JUL-1987 US 071674 PI
CORNELIA MAXIN GOMAN
PC C12N15/09,C12N5/10,C12N15/00,C12N5/00
CC
FH Key Location/Qualifiers
FT source 1..773
/organism="Unidentified".
FEATURES
Location/Qualifiers
1..777
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 188 a 205 c 190 g 194 t
ORIGIN
Query Match 25.4%; Score 127; DB 6; Length 777;
Best Local Similarity 100.0%; Pred. No. 1.4e-65;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 GACATTTCTTCTTGGCAGTACATCTACGTATTAGTCTGCTATTACCATGGTGTATCGGG 76
Db 343 GACATTTCTTCTTGGCAGTACATCTACGTATTAGTCTGCTATTACCATGGTGTATCGGG 402

```

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGAATTTCCAAAGTCTC 136
|||||
Db 403 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGAATTTCCAAAGTCTC 462
|||||
QY 137 CACCCCA 143
|||||
Db 463 CACCCCA 469

RESULT 13
I05430
LOCUS I05430 777 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 16 from Patent EP 0260148.
ACCESSION I05430
VERSION I05430.1 GI:591075

KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 777)
AUTHORS Gorman,C.M.
TITLE Improved recombinant expression method, vector and transformed cells
JOURNAL Patent: EP 0260148-A2 16 16-MAR-1988;
FEATURES Location/Qualifiers
source 1..777
BASE COUNT 188 a 205 c 190 g 194 t
ORIGIN

Query Match 25.4%; Score 127; DB 6; Length 777;
Best Local Similarity 100.0%; Pred. No. 1.4e-65;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 GACTTTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGATGCGG 76
|||||
Db 343 GACTTTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGATGCGG 402
|||||
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGAATTTCCAAAGTCTC 136
|||||
Db 403 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGAATTTCCAAAGTCTC 462
|||||
QY 137 CACCCCA 143
|||||
Db 463 CACCCCA 469

RESULT 14
I08105
LOCUS I08105 777 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 3 from Patent EP 0309237.
ACCESSION I08105
VERSION I08105.1 GI:589184

KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 777)
AUTHORS Gorman,C.M.
TITLE A transient expression system for producing recombinant protein
JOURNAL Patent: EP 0309237-A1 3 29-MAR-1989;
FEATURES Location/Qualifiers
source 1..777
BASE COUNT 188 a 205 c 190 g 194 t
ORIGIN

Query Match 25.4%; Score 127; DB 6; Length 777;
Best Local Similarity 100.0%; Pred. No. 1.4e-65;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 GACTTTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGATGCGG 76
|||||

Db 343 GACTTTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGATGCGG 402
|||||
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGAATTTCCAAAGTCTC 136
|||||
Db 403 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGAATTTCCAAAGTCTC 462
|||||
QY 137 CACCCCA 143
|||||
Db 463 CACCCCA 469

RESULT 15
HS5PIELAB
LOCUS Human cytomegalovirus IE-1 promoter region. 875 bp DNA linear VRL 02-AUG-1993
DEFINITION Human cytomegalovirus IE-1 promoter region.
ACCESSION M64941
VERSION M64941.1 GI:330635

KEYWORDS enhancer region; promoter region.
SOURCE Human herpesvirus 5
ORGANISM Human herpesvirus 5
REFERENCE 1 (bases 1 to 875)
AUTHORS Lehner,R., Stammen,T. and Mach,M.
TITLE Comparative sequence analysis of human cytomegalovirus strains
JOURNAL J. Clin. Microbiol. 29 (11), 2494-2502 (1991)
MEDLINE 92129518
PUBMED 1663509
COMMENT Original source text: Human cytomegalovirus DNA.
Individual isolate: patient G.A.
major immediate early gene enhancer.

FEATURES
source 1..875
Location/Qualifiers
/organism="Human herpesvirus 5"
/mol_type="genomic DNA"
/db_xref="taxon:10959"
/cell_line="infected human foreskin fibroblast"
BASE COUNT 227 a 209 c 201 g 238 t
ORIGIN Hind III F fragment of HCMV strain AD169.

Query Match 25.4%; Score 127; DB 14; Length 875;
Best Local Similarity 100.0%; Pred. No. 1.4e-65;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGATGCGG 76
|||||
Db 503 GACTTTCCTACTTGGCAGTACATCTAGCTATTAGTCATCGCTATTACCATGGTGATGCGG 562
|||||
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGAATTTCCAAAGTCTC 136
|||||
Db 563 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGAATTTCCAAAGTCTC 622
|||||
QY 137 CACCCCA 143
|||||
Db 623 CACCCCA 629

Search completed: February 16, 2004, 01:25:45
Job time : 1992.02 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:

February 15, 2004, 19:07:11 ; Search time 161.871 Seconds
(without alignments)
8338.264 Million cell updates/sec

Title: US-09-921-143-36_COPY_1_500

Perfect score: 500

Sequence: 1 aagttgacattatgcact.....cactgggttggtgacatgcac 500

Scoring table:

OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N Genesec 19Jun03.*

1: /SIDSI/gcgdata/genesec/genesecq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/genesec/genesecq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/genesec/genesecq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/genesec/genesecq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/genesec/genesecq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/genesec/genesecq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/genesec/genesecq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/genesec/genesecq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/genesec/genesecq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/genesec/genesecq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/genesec/genesecq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/genesec/genesecq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/genesec/genesecq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/genesec/genesecq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/genesec/genesecq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/genesec/genesecq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/genesec/genesecq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/genesec/genesecq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/genesec/genesecq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/genesec/genesecq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/genesec/genesecq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/genesec/genesecq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/genesec/genesecq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/genesec/genesecq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/genesec/genesecq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	500	100.0	5283	24	ABK10062
2	127	25.4	281	20	AAZ80727
3	127	25.4	282	20	AAZ80726
4	127	25.4	481	25	ABZ17986
5	127	25.4	500	19	AAV14005
6	127	25.4	589	20	AAI15163
7	127	25.4	590	22	AAH24425
8	127	25.4	599	22	AAF84456
					Expression vector
					Cytomegalovirus-de
					S2 subtraction lib
					Cytomegalovirus (C
					Human cytomegalovi
					Cauliflower mosaic
					Cytomegalovirus (C

9	127	25.4	600	24	ABK16313	Human lung tumour
10	127	25.4	619	22	AAI15664	Cauliflower mosaic
11	127	25.4	738	24	AAI44414	CMV promoter / mou
12	127	25.4	763	21	AAZ95643	Cytomegalovirus de
13	127	25.4	777	9	AAZ82445	Partial expression
14	127	25.4	777	10	AAZ90697	DNA sequence of po
15	127	25.4	884	25	AAZ53783	CMV enhancer/Myelo
16	127	25.4	919	9	AAZ81527	Partial expression
17	127	25.4	919	10	AAZ90695	DNA sequence of po
18	127	25.4	930	7	AAZ60156	Enhancer for eukar
19	127	25.4	930	14	AAZ53550	HCNV (AD169) major
20	127	25.4	930	14	AAZ43524	Sequence of HCMV(A
21	127	25.4	930	18	AAZ77193	HCNV immediate ear
22	127	25.4	987	24	AAZ37261	Adeno-associated v
23	127	25.4	996	20	AAZ25629	Human Sonic hedgeh
24	127	25.4	996	20	AAZ25113	Human Sonic hedgeh
25	127	25.4	996	20	AAZ16193	Human Shh gene PCR
26	127	25.4	996	21	AAZ27888	Human Sonic hedgeh
27	127	25.4	996	21	AAZ30286	Human Sonic hedgeh
28	127	25.4	996	22	AAI66796	Human Shh gene and
29	127	25.4	996	22	AAH28445	Nucleotide sequenc
30	127	25.4	996	24	AAZ23811	Human Sonic hedgeh
31	127	25.4	1070	19	AAV58058	Plasmid CMV-delRT
32	127	25.4	1104	19	AAV03231	DNA encoding CTLA-
33	127	25.4	1215	17	AAZ23597	CMV-PC-PSA promote
34	127	25.4	1216	17	AAZ23595	CMV-PSA promoter.
35	127	25.4	1467	23	ABX04557	DNA encoding Syncr
36	127	25.4	1632	24	AAI44408	CMV promoter / mou
37	127	25.4	1715	24	AAI38382	Chimeric sequence
38	127	25.4	1767	24	AAI38380	Wild-type human AD
39	127	25.4	1767	24	AAI38391	Chimeric sequence
40	127	25.4	1787	24	ABQ55020	Human ovarian anti
41	127	25.4	1813	21	AAZ89475	Transgenic APP DNA
42	127	25.4	1814	21	AAA47118	Reporter gene cons
43	127	25.4	2043	23	ABX04567	DNA encoding N-ter
44	127	25.4	2133	10	AAZ91042	Promoter-enhancer
45	127	25.4	2196	15	AAQ74211	CMVintA promoter a

ALIGNMENTS

RESULT 1

ABK10062
ID ABK10062 standard; DNA; 5283 BP.

XX ABK10062;

AC 21-MAY-2002 (first entry)

XX Expression vector construct pVGI.1 containing VEGF-2 insert.

XX Human; vascular endothelial growth factor 2; VEGF-2; ds; pVGI.1;
XX chronic limb ischaemia; myocardial ischaemia; autoimmune disorder; pHE4a;
XX allergic reaction; organ rejection; inflammatory condition; arrhythmia;
XX hyperproliferative disorder; viral infection; bacterial infection;
XX fungal infection; parasitic infection; cardiovascular disorder; embolism;
XX heart valve disease; aneurysm; arterial occlusive disorder; gene therapy.
XX Synthetic.

XX WO200211769-A1.

XX 14-FEB-2002.

XX 03-AUG-2001; 2001WO-US24658.

XX 04-AUG-2000; 2000US-223276P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Coleman TA;

XX

```
DR WPI; 2002-217153/27.
XX Isolated nucleic acid having expression vector construct with vascular
PT endothelial growth factor-2 insert, useful for treating chronic limb
PT ischaemia or myocardial ischaemia, autoimmune disorders and allergic
PT conditions -
XX PS Disclosure; Fig 31; 241pp; English.
XX CC The invention relates to an isolated nucleic acid comprising pVGI.1
CC expression vector construct containing the vascular endothelial growth
CC factor 2 (VEGF-2) insert. The nucleic acid is useful for producing a host
CC cell by transducing, transforming or transfecting a host cell with the
CC DNA and for treating a patient having chronic limb ischaemia or
CC myocardial ischaemia, or a disease or disorder selected from autoimmune
CC disorders (e.g. rheumatoid arthritis, dermatitis), allergic reactions or
CC conditions (e.g. asthma), organ rejection, inflammatory conditions (e.g.
CC Crohn's disease), hyperproliferative disorders (e.g. Gaucher's disease),
CC diseases due to viral, bacterial, fungal or parasitic infection,
CC cardiovascular disorders (e.g. cardiomyopathy), arrhythmia, heart valve
CC diseases, aneurysms, arterial occlusive disorders and embolism. This
CC sequence represents the pVGI.1 expression vector containing the VEGF-2
CC insert.
XX SQ Sequence 5283 BP; 1244 A; 1422 C; 1417 G; 1200 T; 0 other;
Query Match 100.0%; Score 500; DB 24; Length 5283;
Best Local Similarity 100.0%; Pred. No. 4.1e-262;
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGTTGACCTTATGCGACTTCTCTACTTGGCAGTACATCTAGCTATTAGTCTATCCTAT 60
DB 1 AAGTTGACCTTATGCGACTTCTCTACTTGGCAGTACATCTAGCTATTAGTCTATCCTAT 60
QY 61 TACCATGGTATGCGGTTTGGCAGTACATCAATGGCGTGGATAGCGTTGACTCAG 120
DB 61 TACCATGGTATGCGGTTTGGCAGTACATCAATGGCGTGGATAGCGTTTGAATCAG 120
QY 121 GGGATTTCAGTCTCCACCCCACTGACGTCATGCGAGTTCTTTTGGCACCABAATCA 180
DB 121 GGGATTTCAGTCTCCACCCCACTGACGTCATGCGAGTTCTTTTGGCACCABAATCA 180
QY 181 ACGAGACTTCCAAAATGTGTACAACTCCGCCCACTTACGCAAAATGGCGGTAGGGC 240
DB 181 ACGAGACTTCCAAAATGTGTACAACTCCGCCCACTTACGCAAAATGGCGGTAGGGC 240
QY 241 AACATGCTTATGTAACGGTGATTTAGCAACATGCTTATAGGAGAGAAAACACCGTG 300
DB 241 AACATGCTTATGTAACGGTGATTTAGCAACATGCTTATAGGAGAGAAAACACCGTG 300
QY 301 CATGCCGATTGGTGGAGTAAGGTGGTATGATCGTGTATGATCGTGCCTTGTAGGAAG 360
DB 301 CATGCCGATTGGTGGAGTAAGGTGGTATGATCGTGTATGATCGTGCCTTGTAGGAAG 360
QY 361 GCAACAGACGGGTCTAACAGGATTTGACAGAACCACTGAATTCGCGANTTCAGAGATTT 420
DB 361 GCAACAGACGGGTCTAACAGGATTTGACAGAACCACTGAATTCGCGANTTCAGAGATTT 420
QY 421 GTATTTAAGTCCGAGCTCGATACATAAAGCCGATTTGACCATTCACCACTTGGTGTG 480
DB 421 GTATTTAAGTCCGAGCTCGATACATAAAGCCGATTTGACCATTCACCACTTGGTGTG 480
QY 481 CACCTGGTGGATTCGATC 500
DB 481 CACCTGGTGGATTCGATC 500
RESULT 2
AAX80727
ID AAX80727 standard; DNA; 281 BP.
XX AC AAX80727;
XX
```

```
DT 19-OCT-1999 (first entry)
XX Cytomegalovirus-derived promoter element-2 (-299 to -19).
XX Cytomegalovirus-derived promoter element-2; CMV-derived promoter element;
KW Cytomegalovirus immediate early promoter; human alpha-galactosidase gene;
KW persistent transgene expression; expression cassette; human CFTR gene;
KW human cystic fibrosis transmembrane conductance regulator gene;
KW adenoviral vector; E4 region; transgene; plasmid; ds.
XX Cytomegalovirus.
OS
XX Key Location/Qualifiers
FH promoter 1..281
FT /*tag= a
FT /label= CMV_derived_promoter_element
XX WO9936557-A1.
XX 22-JUL-1999.
XX 15-JAN-1999; 99WO-US00915.
XX 16-JAN-1998; 98US-0071673.
XX (GENZ ) GENZYME CORP.
XX Armentano D, Marshall J, Yew N;
XX WPI; 1999-478911/40.
XX A cytomegalovirus (CMV)-derived promoter element and human albumin
PT gene enhancer elements useful for persistent gene expression
XX Claim 3; Page 49; 56pp; English.
XX The present sequence is a promoter element derived from the
CC cytomegalovirus (CMV) immediate early promoter. The promoter element
CC is used to achieve persistent expression of an operably linked transgene,
CC especially human cystic fibrosis transmembrane conductance regulator
CC gene (CFTR gene) or human alpha-galactosidase gene, in a target cell. The
CC expression cassette comprising the promoter element operably linked to a
CC transgene is delivered to a target cell via a plasmid or an adenoviral
CC vector lacking the E4 region. As the promoter element functions
CC independently of the viral E4 region, this allows the use of vectors
CC containing reduced viral genomes thereby increasing the carrying capacity
CC of the vector while decreasing the potential for host immune reaction.
XX SQ Sequence 281 BP; 67 A; 64 C; 73 G; 77 T; 0 other;
Query Match 25.4%; Score 127; DB 20; Length 281;
Best Local Similarity 100.0%; Pred. No. 8.1e-59;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 GACITTCCTACTTGGCAGTACATCTAGTATTAGTCTATTCGTTATACATGGTATCGGG 76
DB 31 GACITTCCTACTTGGCAGTACATCTAGTATTAGTCTATTCGTTATACATGGTATCGGG 90
QY 77 TTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCACGGGATTTCCAAAGTCTC 136
DB 91 TTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCACGGGATTTCCAAAGTCTC 150
QY 137 CACCCCA 143
DB 151 CACCCCA 157
RESULT 3
AAX80726
ID AAX80726 standard; DNA; 282 BP.
XX AC AAX80726;
XX
```

DT 19-OCT-1999 (first entry)
 XX Cytomegalovirus-derived promoter element-1 (-295 to -14).
 DE
 XX
 KW Cytomegalovirus-derived promoter element-1; CMV-derived promoter element;
 KW Cytomegalovirus immediate early promoter; human alpha-galactosidase gene;
 KW persistent transgene expression; expression cassette; human CFTR gene;
 KW human cystic fibrosis transmembrane conductance regulator gene;
 KW adenoviral vector; E4 region; transgene; plasmid; ds.
 XX
 OS Cytomegalovirus.
 XX
 XX Key Location/Qualifiers
 FH 1-282
 FT promoter /tag= a
 FT /label= CMV_derived_promoter_element
 XX
 XX W09936557-A1.
 XX
 XX 22-JUL-1999.
 XX
 XX 15-JAN-1999; 99WO-US00915.
 XX
 XX 16-JAN-1998; 98US-0071673.
 XX
 XX (GENZ) GENZYME CORP.
 XX
 XX Armentano D, Marshall J, Yew N;
 XX WPI; 1999-478911/40.
 XX
 XX A cytomegalovirus (CMV)-derived promoter element and human albumin
 PT gene enhancer elements useful for persistent gene expression
 PT
 XX
 XX Claim 3; Pages 48-49; 56pp; English.
 XX
 CC The present sequence is a promoter element derived from the
 CC cytomegalovirus (CMV) immediate early promoter. The promoter element
 CC is used to achieve persistent expression of an operably linked transgene,
 CC especially human cystic fibrosis transmembrane conductance regulator
 CC gene (CFTR gene) or human alpha-galactosidase gene, in a target cell. The
 CC expression cassette comprising the promoter element operably linked to a
 CC transgene is delivered to a target cell via a plasmid or an adenoviral
 CC vector lacking the E4 region. As the promoter element functions
 CC independently of the viral E4 region, this allows the use of vectors
 CC containing reduced viral genomes thereby increasing the carrying capacity
 CC of the vector while decreasing the potential for host immune reaction.
 XX
 SQ Sequence 282 BP; 68 A; 65 C; 72 G; 77 T; 0 other;
 Query Match 25.4%; Score 127; DB 20; Length 282;
 Best Local Similarity 100.0%; Pred. No. 8.1e-59;
 Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 GACTTTCCTACTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTATGCGG 76
 Db 27 GACTTTCCTACTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTATGCGG 86
 QY 77 TTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCAGCGGGATTTCGAAGTCTC 136
 Db 87 TTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCAGCGGGATTTCGAAGTCTC 146
 QY 137 CACCCCA 143
 Db 147 CACCCCA 153
 RESULT 4
 AB217986
 ID AB217986 standard; cDNA; 481 BP.
 XX
 AC AB217986;
 XX

DT 23-JAN-2003 (first entry)
 XX S2 subtraction library cancer related clone SEQ ID NO:412.
 DE
 XX
 KW Human; cancer; tumour; therapy; diagnosis; CT antigen; CP antigen;
 KW immune response; virology; immunology; microbiology; molecular biology;
 KW recombinant DNA technology; gene; ss.
 XX
 OS Homo sapiens.
 XX
 XX W0200278516-A2.
 XX
 XX 10-OCT-2002.
 XX
 XX 28-MAR-2002; 2002WO-US10421.
 XX
 XX 30-MAR-2001; 2001US-280255P.
 XX
 XX 28-AUG-2001; 2001US-315563P.
 XX
 XX 09-JAN-2002; 2002US-347313P.
 XX
 XX (CORI-) CORIXA CORP.
 XX
 XX Wang T, Wang S, Bangur CS, Gaiger A;
 XX WPI; 2003-058387/05.
 XX
 XX New immunogenic polynucleotides or polypeptides useful for diagnosing,
 PT preventing and treating cancer expressing CT or CP mRNA antigens, and
 PT in virology, immunology, microbiology, molecular biology and
 PT recombinant DNA techniques -
 XX
 XX Claim 1; SEQ ID 412; 207pp; English.
 XX
 CC ABQ17575 to ABQ20506 represent isolated polynucleotide (I) sequences, and
 CC ABP54446 to ABP54472 represent protein (II) sequences, from the present
 CC invention. (I) and (II) have cytostatic activity and can be used in gene
 CC therapy and vaccines. (I), (II), antibodies and compositions from the
 CC present invention are useful for diagnosing, preventing and treating
 CC cancer, which expresses CT or CP mRNA antigens. They are useful for
 CC stimulating immune response. They can also be useful in virology,
 CC immunology, microbiology, molecular biology and recombinant DNA
 CC techniques.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 481 BP; 121 A; 119 C; 109 G; 132 T; 0 other;
 Query Match 25.4%; Score 127; DB 25; Length 481;
 Best Local Similarity 100.0%; Pred. No. 8e-59;
 Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 GACTTTCCTACTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTATGCGG 76
 Db 315 GACTTTCCTACTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTATGCGG 374
 QY 77 TTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCAGCGGGATTTCGAAGTCTC 136
 Db 375 TTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCAGCGGGATTTCGAAGTCTC 434
 QY 137 CACCCCA 143
 Db 435 CACCCCA 441
 RESULT 5
 AAV14005
 ID AAV14005 standard; cDNA; 500 BP.
 XX
 AC AAV14005;
 XX
 DT 18-JUN-1998 (first entry)
 XX

DE Cytomegalovirus (CMV)-MIE wild-type promoter sequence.

XX Gfi-1; transcription repressor; mutation; cytokine; gene therapy;

KW growth factor independence-1; cytomegalovirus; promoter; ss.

XX Human cytomegalovirus.

OS

XX Key Location/Qualifiers

PH misc_binding /tag= a

FT 331..342 /note= "putative Gfi-1 binding site"

FT 337 /tag= b

FT /note= "G can be substituted by A"

FT 338 /tag= c

FT /note= "A can be substituted by C"

FT 339 /tag= d

FT /note= "T can be substituted by G"

FT 388..399 /tag= e

FT /note= "putative Gfi-1 binding site"

FT 391 /tag= f

FT /note= "A can be substituted by C"

FT 392 /tag= g

FT /note= "T can be substituted by C"

FT 393 /tag= h

FT /note= "C can be substituted by T"

FT CAAT_signal 438..442 /tag= i

FT TATA_signal 471..476 /tag= j

XX WO9748720-A1.

PN

XX 24-DEC-1997.

PD

XX 17-JUN-1997; 97WO-US10486.

PF

XX 17-JUN-1996; 96US-0019808.

PR

XX (FOX-) FOX CHASE CANCER CENT.

PA

XX Grimes HL, Teichlis P, Zweidler-Mckay P;

PI

XX WPI; 1998-063073/06.

DR

XX DNA containing inactive, mutated binding site for Gfi-1

XX transcription repressor - used to increase gene expression in vitro

PT

XX or in vivo, e.g. in gene therapy

PT

XX Disclosure; Fig 3A; 44pp; English.

PS

XX This is the wild-type cytomegalovirus (CMV)-MIE promoter sequence

CC comprising two putative Gfi-1 (growth factor independence-1)

CC transcription repressor binding sites. This promoter can be mutated at

CC the Gfi-1 binding sites to be used in a novel isolated DNA construct

CC which contains at least one mutated binding site for a Gfi-1

CC transcription repressor that hinders or prevents binding of Gfi-1 to

CC this site. The expression vector contains an expression regulatory

CC segment that contains at least one copy of the sequences shown in

CC AAV19671 to AAV19685 linked operably to a coding segment selected from a

CC group of cytokines, interleukins, interferons, growth factors and proto-

CC oncogenes, and an isolated DNA molecule containing one of two 500 base

CC pair sequences shown in AAV14003 and AAV14004. Altering the binding site

CC increases expression of these genes controlled by regulators that

CC include binding sites for Gfi-1 both in cultured cells and in vivo (for

CC gene therapy or DNA vaccination). A vector containing a normal gene

CC under control of a regulator with the mutated binding site can be

CC administered to a patient having a disease associated with an aberrant

XX form of the gene.

XX Sequence 500 BP; 121 A; 126 C; 125 G; 128 T; 0 other;

SQ

Query Match 25.4%; Score 127; DB 19; Length 500;

Best Local Similarity 100.0%; Pred. No. 8e-59;

Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGTATTACCATGTTGATGCGG 76

DB 231 GACTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGTATTACCATGTTGATGCGG 290

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGTTTGACTACGGGGATTTCCAAGTCTC 136

DB 291 TTTTGGCAGTACATCAATGGCGGTGGATAGCGTTTGACTACGGGGATTTCCAAGTCTC 350

QY 137 CACCCCA 143

DB 351 CACCCCA 357

RESULT 6

AAK15163/C

ID AAK15163 standard; DNA; 589 BP.

XX

AC AAK15163;

XX

DT 22-APR-1999 (first entry)

XX

DE Human cytomegalovirus promoter region.

XX

KW apolipoprotein; binding; in vivo transport; nucleic acid; binding domain;

KW nuclear localisation sequence; gene therapy; cancer; cystic fibrosis;

KW non-small cell lung carcinoma; diabetes; arteriosclerosis;

KW Human cytomegalovirus promoter region; ss.

XX

OS Human cytomegalovirus.

XX

PN WO9856938-A1.

XX

PD 17-DEC-1998.

XX

PF 10-JUN-1998; 98WO-US11927.

XX

PR 14-MAY-1998; 98US-0079030.

PR 13-JUN-1997; 97US-0874807.

XX

PA (BAYU) BAYLOR COLLEGE MEDICINE.

XX

PI Guevara JG, Hoogeveen RC, Moore JP;

XX

DR WPI; 1999-070331/06.

XX

PT Composition comprising nucleic acid bound to LDL or VLDL lipoprotein

PT - used for delivering nucleic acid to cells for gene therapy and

PT antisense treatment

XX

XX Example 4; Page 74; 293pp; English.

XX

CC The present sequence represents the Human cytomegalovirus promoter

CC region. The compositions of the invention bind to this sequence.

CC The specification describes a composition that comprises low

CC density lipoprotein (LDL) and apolipoproteins for the binding and

CC in vivo transport of nucleic acids. The composition is used to deliver

CC nucleic acids to eukaryotic cells, in vivo or in vitro, for expressing

CC a therapeutic polypeptide or antisense molecule (or ribozyme).

CC Specifically they are used for gene therapy of cancers (particularly

CC non-small cell lung carcinoma), diabetes, cystic fibrosis and

CC arteriosclerosis.

XX

SQ Sequence 589 BP; 157 A; 140 C; 143 G; 149 T; 0 other;

Query Match 25.4%; Score 127; DB 20; Length 589;
 Best Local Similarity 100.0%; Pred. No. 8e-59; Mismatches 0; Indels 0; Gaps 0;
 Matches 127; Conservative 0;

QY 17 GACTTTCTTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTGATGCGG 76
 DB 276 GACTTTCTTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTGATGCGG 217

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTACACGGGGATTTCCAAAGTCTC 136
 DB 216 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTACACGGGGATTTCCAAAGTCTC 157

QY 137 CACCCCA 143
 DB 156 CACCCCA 150

RESULT 7
 AAH24425
 ID AAH24425 standard; DNA; 590 BP.
 XX
 AC AAH24425;
 XX
 DT 02-AUG-2001 (first entry)
 XX
 DE Cauliflower mosaic virus promoter.
 XX
 DE Cauliflower mosaic virus; promoter; signal peptide; chloroella;
 KW gene expression; protein production; human growth hormone; ds.
 KW
 XX Cauliflower mosaic virus.
 OS
 PN JP2000354490-A.
 XX
 XX 26-DEC-2000.
 XX
 XX 15-JUN-1999; 99JP-0168271.
 XX
 XX 15-JUN-1999; 99JP-0168271.
 XX
 XX (TOYT) TOYOTA JIDOSHA KK.
 XX
 XX WPI; 2001-275809/29.
 XX
 XX New signal peptides useful for the preparation of human growth hormone
 PT and transformed chloroella -
 XX
 XX Example 1; Page 12; 15pp; Japanese.
 PS
 XX The present sequence is provided in a specification relating to signal
 CC peptides for expression and secretion of a protein in chloroella. The
 CC peptides are of the formula:
 CC Met-Ala-Asn-Lys-X₁-1-(Leu)_n-X₂-Ala-Ser-Gly.
 CC X₁ = Ser or Leu;
 CC n₁ = an integer of 5-15;
 CC X₂ = Gly-Ser-Leu or Pro-Leu-Ala.
 CC The signal peptides are useful in the preparation of human growth
 CC hormone and transformed chloroella. Signal peptides, DNA encoding the
 CC peptides, gene expression cassettes, recombinant vectors containing the
 CC cassettes, and transformants having the vectors are provided. The
 CC present sequence is a promoter which may be used in the invention.
 XX
 XX Sequence 590 BP; 149 A; 143 C; 141 G; 157 T; 0 other;
 SQ

Query Match 25.4%; Score 127; DB 22; Length 590;
 Best Local Similarity 100.0%; Pred. No. 8e-59;
 Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCTTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTGATGCGG 76
 DB 314 GACTTTCTTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTGATGCGG 373

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTACACGGGGATTTCCAAAGTCTC 136

Db 374 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTACACGGGGATTTCCAAAGTCTC 433
 QY 137 CACCCCA 143
 DB 434 CACCCCA 440

RESULT 8
 AAF84456
 ID AAF84456 standard; DNA; 599 BP.
 XX
 AC AAF84456;
 XX
 DT 25-JUN-2001 (first entry)
 XX
 DE Cytomegalovirus (CMV) immediate-early (IE) promoter enhancer.
 XX
 KW Cytomegalovirus; CMV; immediate-early; IE promoter enhancer;
 KW pec promoter; chicken beta-actin promoter;
 KW recombinant vector; genetic vaccine; gene therapy; ds.
 XX
 OS Human cytomegalovirus.
 PN JP2001000188-A.
 XX
 XX 09-JAN-2001.
 XX
 XX 22-JUN-1999; 99JP-0174804.
 XX
 XX 22-JUN-1999; 99JP-0174804.
 XX
 XX (JAPG) NIPPON ZEON KK.
 XX
 XX WPI; 2001-285232/30.
 XX
 XX New DNA molecule for use as a promoter for preparing a recombinant
 PT containing the DNA which is used for preparing a vaccine -
 XX
 XX Example 1; Page 12; 15pp; Japanese.
 PS
 XX The invention relates to a 5' and a 3' fragment of the chicken beta-actin
 CC promoter (AAF84452 and AAF84453, respectively), which exhibit promoter
 CC activity. The chicken beta-actin promoter fragments may be used to drive
 CC expression of a heterologous gene in a recombinant vector used as the
 CC active component of a genetic vaccine. The present sequence represents
 CC a cytomegalovirus (CMV) immediate-early (IE) promoter enhancer, a
 CC portion of which was used in the construction of the chimeric pec
 CC promoter (AAF84454) in an exemplification of the invention.
 XX
 XX Sequence 599 BP; 152 A; 144 C; 144 G; 159 T; 0 other;
 SQ

Query Match 25.4%; Score 127; DB 22; Length 599;
 Best Local Similarity 100.0%; Pred. No. 8e-59;
 Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCTTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTGATGCGG 76
 DB 319 GACTTTCTTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTGATGCGG 378

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTACACGGGGATTTCCAAAGTCTC 136
 DB 379 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTACACGGGGATTTCCAAAGTCTC 438

QY 137 CACCCCA 143
 DB 439 CACCCCA 445

RESULT 9
 ABK16313
 ID ABK16313 standard; cDNA; 600 BP.
 XX

AC ABK16313;
 XX 14-MAR-2002 (first entry)
 DT
 XX Human lung tumour protein encoding cDNA #425.
 DE
 XX Human; lung tumour protein; lung cancer; immunostimulant; cytostatic; ss;
 KW PCR primer.
 XX
 OS Homo sapiens.
 XX
 XX WO200202623-A2.
 PN
 XX 10-JAN-2002.
 PD
 XX
 XX 28-JUN-2001; 2001WO-US20975.
 PF
 XX 29-JUN-2000; 2000US-215696P.
 PR
 XX 22-AUG-2000; 2000US-227142P.
 PR
 XX 06-SEP-2000; 2000US-230481P.
 PR
 XX 21-DEC-2000; 2000US-257729P.
 XX
 XX (CORI-) CORIXA CORP.
 PA
 XX Wang T, McNeill PD, Wantanabe Y, Carter D, Henderson RA, Kalos MD;
 PI
 XX WPI; 2002-106602/14.
 DR
 XX Polynucleotides encoding lung tumour polypeptides, useful for treating
 PT lung cancer or stimulating an immune response -
 XX
 PS Claim 1; Page 261-262; 277pp; English.
 CC The invention relates to polynucleotides encoding lung tumour
 CC polypeptides. Compositions comprising the lung tumour polypeptides,
 CC polynucleotides, fusion proteins, antibodies to the polypeptides, T cell
 CC populations, or antigen presenting cells that express the lung tumour
 CC polypeptides, are useful for treating lung cancer or stimulating an
 CC immune response. The polynucleotides and polypeptides are also useful in
 CC the diagnosis and monitoring of lung cancer. Sequences ABK16343
 CC represent cDNA molecules encoding the polypeptides of the invention and
 CC PCR primers used to amplify the cDNA.
 XX
 SQ Sequence 600 BP; 153 A; 152 C; 141 G; 154 T; 0 other;
 Query Match 25.4%; Score 127; DB 24; Length 600;
 Best Local Similarity 100.0%; Pred. No. 8e-59;
 Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCTATCGCTATTACCGTGGTATCGG 76
 DB 278 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCTATCGCTATTACCGTGGTATCGG 337
 QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCGAAGTCTC 136
 DB 338 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCGAAGTCTC 397
 QY 137 CACCCCA 143
 DB 398 CACCCCA 404
 RESULT 10
 AAS15664
 ID AAS15664 standard; DNA; 619 BP.
 XX
 AC AAS15664;
 XX
 XX 29-JAN-2002 (first entry)
 DT
 XX Cauliflower mosaic virus promoter.
 DE
 XX Ecdysone receptor; EcR; ligand binding domain; ds;
 KW

KW retinoid X receptor; RXRalpha; DNA-binding domain; CMV promoter;
 KW transactivation domain; nuclear receptor; ultraspiracle; gene therapy;
 KW protein production; antibody production; high throughput screening;
 KW HTS; transgenic plant; transgenic animal.
 XX
 OS Cauliflower mosaic virus.
 XX
 XX WO200170816-A2.
 PN
 XX 27-SEP-2001.
 PD
 XX
 XX 21-MAR-2001; 2001WO-US09050.
 PF
 XX 22-MAR-2000; 2000US-191355P.
 PR
 XX 20-FEB-2001; 2001US-269799P.
 PR
 XX (ROHM) ROHM & HAAS CO.
 PA
 XX Palli SR, Kapitskaya MZ, Cress DE;
 PI
 XX WPI; 2001-656841/75.
 DR
 XX Ecdysone and retinoid X receptor based inducible gene expression
 PT systems for use in e.g. gene therapy, large scale production of
 PT proteins and cell-based high-throughput screening assays -
 XX
 XX Example 1; Page 140-141; 144pp; English.
 PS
 CC The invention relates to Ecdysone and retinoid X receptor based inducible
 CC gene expression systems useful for modulating gene expression in host
 CC cells. The gene expression system encodes a polypeptide with a
 CC DNA-binding domain recognizes a response element associated with a gene
 CC whose expression is to be modulated and/or a ligand binding domain (LBD)
 CC comprising a LED from a nuclear receptor and a second gene expression
 CC cassette capable of being expressed in a host cell comprising a
 CC polynucleotide sequence encoding a second polypeptide comprising a trans-
 CC activation domain and/or a LBD comprising a LED from a nuclear receptor
 CC other than ultraspiracle (USP) (the trans-activation domain is from a
 CC nuclear receptor other than an ecdysone receptor, a retinoid X receptor
 CC or a USP receptor and the LBDs from the first and second polypeptides are
 CC different and dimerise). The ecdysone and retinoid X receptor based
 CC inducible gene expression systems useful for modulating gene expression
 CC in host cells, for use in gene therapy, large scale production of
 CC proteins and antibodies, cell-based high-throughput screening
 CC assays (HTS), functional genomic and regulation of traits in transgenic
 CC plants and animals. The present sequence represents a cauliflower mosiac
 CC virus, CMV, promoter which is used in an inducible gene expression
 CC system of the invention.
 XX

SQ Sequence 619 BP; 150 A; 164 C; 155 G; 150 T; 0 other;
 Query Match 25.4%; Score 127; DB 22; Length 619;
 Best Local Similarity 100.0%; Pred. No. 8e-59;
 Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCTATCGCTATTACCGTGGTATCGG 76
 DB 254 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCTATCGCTATTACCGTGGTATCGG 313
 QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCGAAGTCTC 136
 DB 314 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCGAAGTCTC 373
 QY 137 CACCCCA 143
 DB 374 CACCCCA 380

RESULT 11
 AAL44414
 ID AAL44414 standard; DNA; 738 BP.
 XX
 AC AAL44414;

```
XX 31-OCT-2002 (first entry)
XX CMV promoter / mouse G-CSF receptor (G-CSFR) gene 9 - chimeric sequenc.
XX Mouse; murine; ds; gene of interest isolation; CMV; gene; promoter;
XX novel gene discovery; novel gene isolation; large library production;
XX G-CSF receptor; G-CSFR; chimeric.
XX Chimeric - Mus sp.
XX Chimeric - Cytomegalovirus.
XX WO200236760-A2.
XX 10-MAY-2002.
XX 30-OCT-2001; 2001WO-US45045.
XX 30-OCT-2000; 2000US-0702177.
XX (NOVO) NOVOZYMES BIOTECH INC.
XX (ZYMO) ZYMOGENETICS.
XX Lin J, Yaver D, Foster D, Holly R;
XX WPI; 2002-566503/60.
XX Production of a mutant mammalian cell for isolating genes, comprising
XX integrating a nucleic acid construct into the cell genome at sites
XX generated by the restriction enzyme, and selecting cell with trait of
XX interest.
XX Example 11; Fig 14; 106pp; English.
XX The invention comprises a method for producing a mutant mammalian cell.
XX The method involves introducing a restriction enzyme and a nucleic acid
XX regulatory sequence into mammalian cells - for integrating the nucleic
XX acid construct into the mammalian cell genome at sites generated by the
XX restriction enzyme. Mutant mammalian cells having a trait of interest can
XX then be selected. The method of the invention is useful for isolating a
XX gene controlling a trait of interest from a mammalian cell. The method is
XX useful for discovering and isolating new genes. The method of the
XX invention can be used to create large libraries of mammalian cells which
XX have a low transfection efficiency. The method of the invention is also
XX suitable for over-expressing a known endogenous gene that is expressed
XX poorly. The present DNA sequence represents a chimeric gene sequence
XX comprising a Cytomegalovirus (CMV) promoter and a murine G-CSF receptor
XX (G-CSFR) gene.
XX Sequence 738 BP; 188 A; 178 C; 168 G; 204 T; 0 other;
XX
XX Query Match 25.4%; Score 127; DB 24; Length 738;
XX Best Local Similarity 100.0%; Pred. No. 8e-59;
XX Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 17 GACTTTCCTACTTGGCAGTACATCATCGTATTAGTCATCGCTATTACCATGTCATGCGG 76
XX 335 GACTTTCCTACTTGGCAGTACATCATCGTATTAGTCATCGCTATTACCATGTCATGCGG 394
XX
XX QY 77 TTTTGGCAGTACATCAATGGGCGTGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 136
XX 395 TTTTGGCAGTACATCAATGGGCGTGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 454
XX
XX QY 137 CACCCCA 143
XX 455 CACCCCA 461
XX
XX RESULT 12
XX AAZ95643
XX ID AAZ95643 standard; DNA; 763 BP.
XX AC AAZ95643;
```

```
XX 08-JUN-2000 (first entry)
XX Cytomegalovirus derived promoter sequence SEQ ID NO:1.
XX
XX Cytomegalovirus; CMV; promoter; human growth hormone; PGH; ovine;
XX prostaglandin G/H synthase; plasmid; prostanoic; circular;
XX cardiant; thrombolytic; antiinflammatory; gene therapy;
XX pulmonary embolus; myocardial infarction; lung disease;
XX adult respiratory distress syndrome; ds.
XX Human cytomegalovirus.
XX OS
XX US6030638-A.
XX 29-FEB-2000.
XX 02-JUN-1995; 95US-0459493.
XX 19-AUG-1991; 91US-0746941.
XX 21-JUN-1993; 93US-0080221.
XX (UYVA-) UNIV VANDERBILT.
XX Meyrick B, Canonico A, Brigham K, Conary JT;
XX WPI; 2000-205193/18.
XX Increasing prostanoic production in vivo comprising delivering and
XX hyperexpressing a prostaglandin synthase gene in cells, useful for
XX treating e.g. myocardial infarction and general lung disease -
XX Disclosure; Column 7-10; 11pp; English.
XX
XX A method has been developed of increasing prostanoic production in vivo
XX comprising delivering and hyperexpressing a prostaglandin synthase gene
XX in cells. The method can be used for gene therapy. The method is useful
XX for the treatment of diseases such as pulmonary embolus, myocardial
XX infarction and general lung disease (e.g. adult respiratory distress
XX syndrome). Generally the present invention provides a plasmid comprising
XX a cytomegalovirus (CMV) derived promoter sequence driving the coding
XX region for ovine prostaglandin G/H synthase (PGH). In order to increase
XX the expression of the gene, the construct also contains a short
XX translation augmenting sequence and a portion of the 3' untranslated
XX region from the human growth hormone (hGH) gene. The present sequence
XX represents the CMV promoter sequence used in the exemplification of the
XX present invention.
XX Sequence 763 BP; 203 A; 174 C; 164 G; 222 T; 0 other;
XX
XX Query Match 25.4%; Score 127; DB 21; Length 763;
XX Best Local Similarity 100.0%; Pred. No. 8e-59;
XX Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 17 GACTTTCCTACTTGGCAGTACATCATCGTATTAGTCATCGCTATTACCATGTCATGCGG 76
XX 492 GACTTTCCTACTTGGCAGTACATCATCGTATTAGTCATCGCTATTACCATGTCATGCGG 551
XX
XX QY 77 TTTTGGCAGTACATCAATGGGCGTGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 136
XX 552 TTTTGGCAGTACATCAATGGGCGTGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 611
XX
XX QY 137 CACCCCA 143
XX 612 CACCCCA 618
XX
XX RESULT 13
XX AA82445
XX ID AA82445 standard; DNA; 777 BP.
XX AC AA82445;
```

```
DT 06-MAR-1992 (first entry)
XX Partial expression vector pF8C8SS sequence.
DE Expression vector; promoter; heterologous protein; mammalian vector;
KW ds.
XX Cytomegalovirus.
OS Key Location/Qualifiers
FH misc_feature 2 /*tag= a
FT /note= "from pPMLCMV beginning to HindIII,
FT enhancers and promoter"
FT 612
FT /*tag= b
FT /note= "starts here"
PN EP260148-A.
XX 16-MAR-1988.
PD 11-SEP-1987; 87EP-0308060.
XX 09-JUL-1987; 87US-0071674.
PR 12-SEP-1986; 86US-0907185.
PR 25-SEP-1987; 87US-0101712.
XX (GETH ) GENENTECH INC.
PA Gorman CM;
XX WPI; 1988-072583/11.
XX Continuous expression of heterologous protein in eukaryotic cells
PT - by transfecting with vector contg. stabilising sequence,
PT promoter, structural gene, poly:adenylate sequence and
PT termination site
XX Example; Fig 19; 48pp; English.
XX In the expression vectors of the invention, the promoter is
CC specifically from the immediate early gene of human cytomegalovirus
CC (CMV) or from simian virus (SV40). The stabilizing SQ contains 1 or
CC 2 splice donor-intron-acceptor SQs, esp. from the specified CMV gene,
CC or the intron is from CMV or the Ig variable region, while the
CC splice-acceptor component corresponds to the Ig acceptor SQ. The
CC vector may also include (upstream of the promoter) an enhancer, esp.
CC also from SV40.
XX Sequence 777 BP; 188 A; 205 C; 190 G; 194 T; 0 other;
SQ Query Match 25.4%; Score 127; DB 9; Length 777;
Best Local Similarity 100.0%; Pred. No. 8e-59;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 GACTTCTCTACTTGCGAGTACATCTATAGTCATCGCTATACCATGTCATCGCG 76
DB 343 GACTTCTCTACTTGCGAGTACATCTATAGTCATCGCTATACCATGTCATCGCG 402
QY 77 TTTTGGCAGTACATCAATGGCGGTGATAGCGTTTGACTCACGGGATTTCCAGTCTC 136
DB 403 TTTTGGCAGTACATCAATGGCGGTGATAGCGTTTGACTCACGGGATTTCCAGTCTC 462
QY 137 CACCCCA 143
DB 463 CACCCCA 469
RESULT 14
AAN90697
ID AAN90697 standard; DNA; 777 BP.
XX
```

```
AC AAN90697;
XX 25-MAR-2003 (updated)
DT 15-JUL-1990 (first entry)
XX DNA sequence of portion of plasmid pF8C8SS containing the cytomegalovirus
DE enhancer and promoter and an engineered stabilising.
XX Plasmid pF8C8SS; cytomegalovirus; SV40; enhancer; promoter;
KW stabilising sequence; splice donor intron sequence; Ig region;
KW splice acceptor sequence; human embryonic kidney cells (293); JW2.
XX Cytomegalovirus; SV40.
OS Key Location/Qualifiers
FH promoter 1..732
FT /*tag= a
FT /note="cytomegalovirus enhancer, promoter and leader"
FT misc_feature 733..736
FT /*tag= b
FT /note="stabilising sequence including the engineered
FT splice donor and acceptor sequence"
FT 737..777
FT /*tag= c
FT /note="remaining leader"
FT misc_RNA 612
FT /*tag= d
FT /note="Begin RNA"
XX EP309237-A.
PN 29-MAR-1989.
XX 22-SEP-1988; 89EP-0308784.
PR 25-SEP-1987; 87US-0101712.
XX (GETH ) GENENTECH INC.
PA Gorman CM;
XX WPI; 1989-095394/13.
XX Transient expression system for recombinant proteins -
PT comprising eukaryotic host cell transfected with vector
PT encoding trans-activating protein and expression vector
XX Disclosure; 44 pp.; English.
XX Expression vector pF8C8SS contains the cytomegalovirus enhancer and
CC promoter, an engineered stabilising sequence, the cDNA encoding factor
CC VIII and the SV40 polyadenylation site. The entire intron region
CC an engineered donor and acceptor sequence was deleted and replaced by
CC an engineered stabilising sequence, pF8C8SS is used in a method for
CC prodn. of a desired heterologous protein in a eukaryotic host cell.
CC The method comprises: transfecting a eukaryotic host cell with a vector
CC encoding trans-activating protein; transfecting the host cell with
CC an expression vector having a sequence of a ds DNA comprising, a
CC stabilising sequence downstream of a promoter and upstream of a DNA
CC -adenylation sequence upstream to a transcription termination site.
CC The pref. host cells are human embryonic kidney cells (293) or JW2.
CC The method provides useful quantities of a desired protein in a
CC relatively short period of time without having to establish continuous
CC prodn. The transient expression system optimises the interaction between
CC specific vector components and certain trans-activating proteins.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 777 BP; 188 A; 205 C; 190 G; 194 T; 0 other;
SQ Query Match 25.4%; Score 127; DB 10; Length 777;
Best Local Similarity 100.0%; Pred. No. 8e-59;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 17 GACTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGTATTACCATGGTGATCGCG 76
 |||||
 Db 343 GACTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGTATTACCATGGTGATCGCG 402
 |||||
 QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGATTTCCAAAGTCTC 136
 |||||
 Db 403 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGATTTCCAAAGTCTC 462
 |||||
 QY 137 CACCCCA 143
 |||||
 Db 463 CACCCCA 469
 |||||

RESULT 15

AAD53783

ID AAD53783 standard; DNA; 884 BP.

XX

AC AAD53783;

XX 28-MAY-2003 (first entry)

DT

XX CMV enhancer/Myeloproliferative sarcoma virus LTR promoter construct DNA.

DE

KW Transmembrane activator; calcium modulator; nephrotropic; antibacterial;
 TACI; tumour necrosis factor-like protein; ZTNF2; ZTNF4; immunoglobulin;
 anaemia; gene therapy; cytostatic; antiinflammatory; immunosuppressive;
 glomerulonephritis; asthma; bronchitis; graft rejection; septic shock;
 dermatological; neuroprotective; cyclophilin ligand-interactor;
 autoimmune disease; systemic lupus erythematosus; multiple sclerosis;
 diabetes mellitus; rheumatoid arthritis; renal disease; inflammation;
 Cytomegalovirus; CMV; chimeric; gene; ds.

XX

OS Chimeric - Cytomegalovirus.

XX Chimeric - Myeloproliferative sarcoma virus.

XX

FH Key Location/Qualifiers

FT misc_feature 1..407

FT /tag= a

FT /note= "CMV enhancer"

FT misc_feature 408..884

FT /tag= b

FT /note= "Myeloproliferative sarcoma virus LTR promoter"

XX

PN WO200294852-A2.

XX

PD 28-NOV-2002.

XX

PF 20-MAY-2002; 2002WO-US15910.

XX

PR 24-MAY-2001; 2001US-293343P.

XX

PA (ZYMO) ZYMOGENETICS INC.

XX

PI Rixon MW, Gross JA;

XX

XX WPI; 2003-148455/14.

XX

Transmembrane activator and calcium modulator and cyclophilin
 ligand-interactor (TACI)-immunoglobulin fusion protein, for treating
 cancer or diabetes, comprises a TACI receptor group and an
 immunoglobulin group -

XX

PS Disclosure; Column 141; 71pp; English.

XX

CC The invention relates to fusion proteins comprising transmembrane
 activator and calcium modulator and cyclophilin ligand-interactor (TACI)
 receptor group that binds tumour necrosis factor-like protein (ZTNF)2 or
 ZTNF4; and an immunoglobulin group comprising a constant region of an
 immunoglobulin. The invention is used to manufacture a medicament for
 inhibiting the proliferation of tumour cells in a mammalian subject.
 CC The composition comprising the fusion protein may also be used in
 treating autoimmune diseases (e.g. systemic lupus erythematosus,

CC multiple sclerosis, diabetes mellitus, rheumatoid arthritis and asthma),
 CC renal diseases (e.g. glomerulonephritis), bronchitis, inflammation,
 CC graft rejection, anaemia and septic shock. The fusion proteins are
 CC also used in gene therapy. The present sequence is Cytomegalovirus (CMV)
 CC enhancer/Myeloproliferative sarcoma virus LTR promoter construct DNA
 CC used in the invention.

XX Sequence 884 BP; 229 A; 236 C; 214 G; 205 T; 0 other;

Query Match 25.4%; Score 127; DB 25; Length 884;

Best Local Similarity 100.0%; Pred. No. 7,9e-59;

Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGTATTACCATGGTGATCGCG 76

Db 256 GACTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGTATTACCATGGTGATCGCG 315

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGATTTCCAAAGTCTC 136

Db 316 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGATTTCCAAAGTCTC 375

QY 137 CACCCCA 143

Db 376 CACCCCA 382

Search completed: February 15, 2004, 22:40:23

Job time : 165.871 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 10:22:55 ; Search time 36.7706 Seconds
(without alignments)
6001.851 Million cell updates/sec

Title: US-09-921-143-36_COPY_1_500

Perfect score: 500
Sequence: 1 aagcttgacctatgcgact.....caoctgggttgatcgatc 500

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 569978 seqs, 220691566 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA*

- 1: /cgn2_6/prodata/2/ina/5A_COMB.seq*
- 2: /cgn2_6/prodata/2/ina/5B_COMB.seq*
- 3: /cgn2_6/prodata/2/ina/6A_COMB.seq*
- 4: /cgn2_6/prodata/2/ina/6B_COMB.seq*
- 5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq*
- 6: /cgn2_6/prodata/2/ina/backfiles1.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	127	25.4	753	US-08-459-493-1	Sequence 1, Appli
2	127	25.4	930	US-08-029-022-2	Sequence 2, Appli
3	127	25.4	930	US-08-029-022-4	Sequence 4, Appli
4	127	25.4	930	US-08-246-376-2	Sequence 2, Appli
5	127	25.4	930	US-08-246-376-4	Sequence 4, Appli
6	127	25.4	930	US-07-972-135-2	Sequence 2, Appli
7	127	25.4	930	US-07-972-135-4	Sequence 4, Appli
8	127	25.4	930	US-08-256-004-2	Sequence 2, Appli
9	127	25.4	930	US-09-006-841-2	Sequence 2, Appli
10	127	25.4	930	US-09-006-841-4	Sequence 4, Appli
11	127	25.4	930	PCT-US93-05366-2	Sequence 2, Appli
12	127	25.4	930	PCT-US93-05366-4	Sequence 4, Appli
13	127	25.4	1078	US-09-310-842-1	Sequence 1, Appli
14	127	25.4	1215	US-08-522-841-7	Sequence 7, Appli
15	127	25.4	1216	US-08-522-841-3	Sequence 3, Appli
16	127	25.4	1318	US-09-310-842-3	Sequence 3, Appli
17	127	25.4	1417	US-09-310-842-5	Sequence 5, Appli
18	127	25.4	1467	US-09-800-170-27	Sequence 27, Appli
19	127	25.4	1645	US-09-310-842-2	Sequence 2, Appli
20	127	25.4	1870	US-09-310-842-4	Sequence 4, Appli
21	127	25.4	2043	US-09-800-170-47	Sequence 47, Appli
22	127	25.4	3125	US-08-037-816A-13	Sequence 13, Appli
23	127	25.4	3125	US-08-530-146-13	Sequence 13, Appli
24	127	25.4	3547	US-09-340-798A-43	Sequence 43, Appli
25	127	25.4	3796	US-09-470-661A-32	Sequence 32, Appli
26	127	25.4	3853	US-08-801-092-5	Sequence 5, Appli
27	127	25.4	3853	US-09-315-113-5	Sequence 5, Appli

28 127 25.4 3925 3 US-09-011-745-9 Sequence 9, Appli
29 127 25.4 3944 1 US-07-678-408A-1 Sequence 1, Appli
30 127 25.4 3987 4 US-03-082-649B-83 Sequence 83, Appli
31 127 25.4 3987 4 US-09-082-649B-84 Sequence 84, Appli
32 127 25.4 4026 3 US-08-801-092-19 Sequence 19, Appli
33 127 25.4 4026 4 US-09-315-113-19 Sequence 19, Appli
34 127 25.4 4059 2 US-08-252-493C-4 Sequence 4, Appli
35 127 25.4 4059 3 US-09-276-197-4 Sequence 4, Appli
36 127 25.4 4059 4 US-08-487-283A-3 Sequence 3, Appli
37 127 25.4 4059 5 PCT-US95-0755A-4 Sequence 4, Appli
38 127 25.4 4059 5 PCT-US96-0561A-11 Sequence 11, Appli
39 127 25.4 4069 4 US-09-170-496D-287 Sequence 287, App
40 127 25.4 4069 4 US-09-170-496D-288 Sequence 288, App
41 127 25.4 4072 4 US-09-770-315-4 Sequence 4, Appli
42 127 25.4 4249 3 US-08-801-092-33 Sequence 33, Appli
43 127 25.4 4249 4 US-09-315-113-33 Sequence 33, Appli
44 127 25.4 4278 4 US-09-503-799-2 Sequence 2, Appli
45 127 25.4 4283 1 US-08-343-401A-3 Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-459-493-1
; Sequence 1, Application US/08459493
; Patent No. 6030638
; GENERAL INFORMATION:
; APPLICANT: Brigham, Kenneth
; APPLICANT: Conary, Jon T.
; APPLICANT: Canonico, Angelo
; APPLICANT: Meyrick, Barbara
; TITLE OF INVENTION: PLASMID FOR IN VIVO EXPRESSION
; TITLE OF INVENTION: PROTAGLANDIN SYNTHASE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Tilton, Fallon, Lungmus & Chestnut
; STREET: 100 S. Wacker Drive - Suite 960
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60606-4002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/459,493
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/080,221
; FILING DATE: 21-JUNE-1993
; APPLICATION NUMBER: US 07/746,941
; FILING DATE: 19-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: VU9115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (313) 456-8000
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 763 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: promoter

LOCATION: 1..763
US-08-459-493-1

Query Match 25.4%; Score 127; DB 3; Length 763;
Best Local Similarity 100.0%; Pred. No. 2.5e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGCTGATCGG 76
DB 492 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGCTGATCGG 551

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCGAAGTCTC 136
DB 552 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCGAAGTCTC 611

QY 137 CACCCCA 143
DB 612 CACCCCA 618

RESULT 2
US-08-029-022-2
; Sequence 2, Application US/08029022
; Patent No. 5641662
; GENERAL INFORMATION:
; APPLICANT: Debs, Robert J.
; APPLICANT: Zhu, Ning
; TITLE OF INVENTION: TRANSFECTION OF LUNG VIA AEROSOLIZED
; TITLE OF INVENTION: TRANSGENE DELIVERY
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/029,022
; FILING DATE: 19930310
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 809,291
; FILING DATE: 17 DEC 91
; APPLICATION NUMBER: 972,135
; FILING DATE: 05 NOV 92
; APPLICATION NUMBER: PCT/US92/11008
; FILING DATE: 17 DEC 92
; ATTORNEY/AGENT INFORMATION:
; NAME: Rae-Venter, Barbara
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: 05935/008001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 854-5277
; TELEFAX: (415) 854-0875
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 930
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-029-022-2

Query Match 25.4%; Score 127; DB 1; Length 930;
Best Local Similarity 100.0%; Pred. No. 2.5e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGCTGATCGG 76
DB 469 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGCTGATCGG 528

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCGAAGTCTC 136
DB 529 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCGAAGTCTC 588

QY 137 CACCCCA 143
DB 589 CACCCCA 595

RESULT 3
US-08-029-022-4
; Sequence 4, Application US/08029022
; Patent No. 5641662
; GENERAL INFORMATION:
; APPLICANT: Debs, Robert J.
; APPLICANT: Zhu, Ning
; TITLE OF INVENTION: TRANSFECTION OF LUNG VIA AEROSOLIZED
; TITLE OF INVENTION: TRANSGENE DELIVERY
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/029,022
; FILING DATE: 19930310
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 809,291
; FILING DATE: 17 DEC 91
; APPLICATION NUMBER: 972,135
; FILING DATE: 05 NOV 92
; APPLICATION NUMBER: PCT/US92/11008
; FILING DATE: 17 DEC 92
; ATTORNEY/AGENT INFORMATION:
; NAME: Rae-Venter, Barbara
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: 05935/008001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 854-5277
; TELEFAX: (415) 854-0875
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 930
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-029-022-4

Query Match 25.4%; Score 127; DB 1; Length 930;
Best Local Similarity 100.0%; Pred. No. 2.5e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGCTGATCGG 76
DB 469 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGCTGATCGG 528

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCGAAGTCTC 136
DB 529 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCGAAGTCTC 588

QY 137 CACCCCA 143
Db 589 CACCCCA 595

RESULT 4

US-08-246-376-2
; Sequence 2, Application US/08246376
; Patent No. 5827703
; GENERAL INFORMATION:
; APPLICANT: Robert J. Debs
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR IN VIVO
; TITLE OF INVENTION: GENE THERAPY
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94025

COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/246,376
; FILING DATE:

CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/992,687
; FILING DATE: December 17, 1992
; APPLICATION NUMBER: 07/927,200
; FILING DATE: August 6, 1992
; APPLICATION NUMBER: 07/894,498
; FILING DATE: June 4, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Barbara Rae-Venter
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: 05935/005US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 854-5277
; TELEX: 200154

INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 930
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-246-376-2

Query Match 25.4%; Score 127; DB 1; Length 930;
Best Local Similarity 100.0%; Pred. No. 2.5e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGG 76
Db 469 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGG 528
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGAATTCCAAGTCTC 136
Db 529 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGAATTCCAAGTCTC 588
QY 137 CACCCCA 143
Db 589 CACCCCA 595

RESULT 5

US-08-246-376-4
; Sequence 4, Application US/08246376
; Patent No. 5827703
; GENERAL INFORMATION:
; APPLICANT: Robert J. Debs
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR IN VIVO
; TITLE OF INVENTION: GENE THERAPY
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94025

COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/246,376
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/992,687
; FILING DATE: December 17, 1992
; APPLICATION NUMBER: 07/927,200
; FILING DATE: August 6, 1992
; APPLICATION NUMBER: 07/894,498
; FILING DATE: June 4, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Barbara Rae-Venter
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: 05935/005US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 854-5277
; TELEX: 200154

INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 930
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-246-376-4

Query Match 25.4%; Score 127; DB 1; Length 930;
Best Local Similarity 100.0%; Pred. No. 2.5e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGG 76
Db 469 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGG 528
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGAATTCCAAGTCTC 136
Db 529 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGAATTCCAAGTCTC 588
QY 137 CACCCCA 143
Db 589 CACCCCA 595

RESULT 6

US-07-972-135-2
; Sequence 2, Application US/07972135
; Patent No. 5858784
; GENERAL INFORMATION:
; APPLICANT: Robert J. Debs
; TITLE OF INVENTION: EXPRESSION OF CLONED
; TITLE OF INVENTION: GENES IN THE

```
;; TITLE OF INVENTION: LUNG BY AEROSOL - AND
;; TITLE OF INVENTION: LIPOSOME-BASED
;; TITLE OF INVENTION: DELIVERY
;; NUMBER OF SEQUENCES: 4
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Fish & Richardson
;; STREET: 2200 Sand Hill Road,
;; CITY: Menlo Park
;; STATE: California
;; COUNTRY: U.S.A.
;; ZIP: 94025
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;; COMPUTER: IBM PS/2 Model 502 or
;; COMPUTER: 55SX
;; OPERATING SYSTEM: MS-DOS (Version 5.0)
;; SOFTWARE: WordPerfect (Version
;; SOFTWARE: 5.1)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/07/972,135
;; FILING DATE: No. 5858784ember 5, 1992
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 07/809,291
;; FILING DATE: December 17, 1991
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Barbara Rae-Venter
;; REGISTRATION NUMBER: 32,750
;; REFERENCE/DOCKET NUMBER: 05935/007US1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 854-5277
;; TELEFAX: (415) 854-0875
;; TELEX: 200154
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 930
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-07-972-135-2

Query Match 25.4%; Score 127; DB 2; Length 930;
Best Local Similarity 100.0%; Pred. No. 2.5e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 GAC TTCTCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGGG 76
Db 469 GAC TTCTCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGGG 528

Qy 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 136
Db 529 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 588

Qy 137 CACCCCA 143
Db 589 CACCCCA 595

RESULT 7
US-07-972-135-4
; Sequence 4, Application US/07972135
; Patent No. 5858784
; GENERAL INFORMATION:
; APPLICANT: Robert J. Debs
; TITLE OF INVENTION: EXPRESSION OF CLONED
; TITLE OF INVENTION: GENES IN THE
; TITLE OF INVENTION: LUNG BY AEROSOL - AND
; TITLE OF INVENTION: LIPOSOME-BASED
; TITLE OF INVENTION: DELIVERY
; NUMBER OF SEQUENCES: 4
```

```
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Fish & Richardson
;; STREET: 2200 Sand Hill Road,
;; CITY: Menlo Park
;; STATE: California
;; COUNTRY: U.S.A.
;; ZIP: 94025
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;; COMPUTER: IBM PS/2 Model 502 or
;; COMPUTER: 55SX
;; OPERATING SYSTEM: MS-DOS (Version 5.0)
;; SOFTWARE: WordPerfect (Version
;; SOFTWARE: 5.1)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/07/972,135
;; FILING DATE: No. 5858784ember 5, 1992
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 07/809,291
;; FILING DATE: December 17, 1991
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Barbara Rae-Venter
;; REGISTRATION NUMBER: 32,750
;; REFERENCE/DOCKET NUMBER: 05935/007US1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 854-5277
;; TELEFAX: (415) 854-0875
;; TELEX: 200154
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 930
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-07-972-135-4

Query Match 25.4%; Score 127; DB 2; Length 930;
Best Local Similarity 100.0%; Pred. No. 2.5e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 GAC TTCTCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGGG 76
Db 469 GAC TTCTCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGGG 528

Qy 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 136
Db 529 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 588

Qy 137 CACCCCA 143
Db 589 CACCCCA 595

RESULT 8
US-08-256-004-2
; Sequence 2, Application US/08256004
; Patent No. 6001644
; GENERAL INFORMATION:
; APPLICANT: Robert J. Debs
; APPLICANT: Ning Zhu
; TITLE OF INVENTION: IN VIVO TRANSFECTION WITH A CFTR CODING
; TITLE OF INVENTION: SEQUENCE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
; STREET: 5 Palo Alto Square
; CITY: Palo Alto
; STATE: California
; COUNTRY: U.S.A.
```

ZIP: 94306-2155
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/256,004
FILING DATE: August 22, 1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/11004
FILING DATE: December 17, 1992
APPLICATION NUMBER: 07/972,135
FILING DATE: NO. 6001644ember 5, 1992
APPLICATION NUMBER: 07/927,200,
FILING DATE: August 6, 1992
APPLICATION NUMBER: 07/894,498
FILING DATE: June 4, 1992
APPLICATION NUMBER: 07/809,291
FILING DATE: December 17, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Neeley, Richard L.
REGISTRATION NUMBER: 30,092
REFERENCE/DOCKET NUMBER: UCSF-008/00US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 843-5070
TELEFAX: (415) 857-0663
TELEX: 380816COOLEYPA
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 930
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-256-004-2

Query Match 25.4%; Score 127; DB 3; Length 930;
Best Local Similarity 100.0%; Pred. No. 2.5e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACCTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGTATTACCATGGTGATCGG 76
Db 469 GACCTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGTATTACCATGGTGATCGG 528

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGGGGATTTCCAAAGTCTC 136
Db 529 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGGGGATTTCCAAAGTCTC 588

QY 137 CACCCCA 143
Db 589 CACCCCA 595

RESULT 9
US-09-006-841-2
Sequence 2, Application US/09006841
Patent No. 6468798
GENERAL INFORMATION:
APPLICANT: Robert J. Debs
TITLE OF INVENTION: EXPRESSION OF CLONED
GENES IN THE
TITLE OF INVENTION: LUNG BY AEROSOL - AND
TITLE OF INVENTION: LIPOSOME-BASED
DELIVERY
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 2200 Sand Hill Road,
SUITE: Suite 100
CITY: Menlo Park
STATE: California
COUNTRY: U.S.A.

ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/006,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/972,135
FILING DATE: No. 6468798ember 5, 1992
APPLICATION NUMBER: 07/809,291
FILING DATE: December 17, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Barbara Rae-Venter
REGISTRATION NUMBER: 32,750
REFERENCE/DOCKET NUMBER: 05935/007U51
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 854-5277
TELEFAX: (415) 854-0875
TELEX: 200154
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 930
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-006-841-2

Query Match 25.4%; Score 127; DB 4; Length 930;
Best Local Similarity 100.0%; Pred. No. 2.5e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACCTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGTATTACCATGGTGATCGG 76
Db 469 GACCTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGTATTACCATGGTGATCGG 528

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGGGGATTTCCAAAGTCTC 136
Db 529 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGGGGATTTCCAAAGTCTC 588

QY 137 CACCCCA 143
Db 589 CACCCCA 595

RESULT 10
US-09-006-841-4
Sequence 4, Application US/09006841
Patent No. 6468798
GENERAL INFORMATION:
APPLICANT: Robert J. Debs
TITLE OF INVENTION: EXPRESSION OF CLONED
GENES IN THE
TITLE OF INVENTION: LUNG BY AEROSOL - AND
TITLE OF INVENTION: LIPOSOME-BASED
DELIVERY
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 2200 Sand Hill Road,
SUITE: Suite 100
CITY: Menlo Park
STATE: California
COUNTRY: U.S.A.

ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/006,841
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/07/972,135
;; FILING DATE: No. 6468798ember 5, 1992
;; APPLICATION NUMBER: 07/809,291
;; FILING DATE: December 17, 1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Barbara Rae-Venter
;; REGISTRATION NUMBER: 32,750
;; REFERENCE/DOCKET NUMBER: 05935/007US1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 854-5277
;; TELEFAX: (415) 854-0875
;; TELEX: 200154
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 930
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-09-006-841-4

Query Match 25.4%; Score 127; DB 4; Length 930;
Best Local Similarity 100.0%; Pred. No. 2.5e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTTATTAGTCATCGCTATTACCATGGTGATCGGG 76
Db 469 GACTTTCCTACTTGGCAGTACATCTACGTTATTAGTCATCGCTATTACCATGGTGATCGGG 528

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTACCGGGATTTCCTCAAGTCTC 136
Db 529 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTACCGGGATTTCCTCAAGTCTC 588

QY 137 CACCCCA 143
Db 589 CACCCCA 595

RESULT 11
PCT-US93-05366-2
; Sequence 2, Application PC/TUS9305366
; GENERAL INFORMATION:
; APPLICANT: Robert J. Debs
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR IN VIVO
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05366
; FILING DATE: 04-JUN-1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/992,687
; FILING DATE: December 17, 1992
; APPLICATION NUMBER: 07/927,200
; FILING DATE: August 6, 1992
; APPLICATION NUMBER: 07/894,498
; FILING DATE: June 4, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Barbara Rae-Venter
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: 05935/005US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 854-5277
; TELEFAX: (415) 854-0875

;; FILING DATE: June 4, 1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Barbara Rae-Venter
;; REGISTRATION NUMBER: 32,750
;; REFERENCE/DOCKET NUMBER: 05935/005US1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 854-5277
;; TELEFAX: (415) 854-0875
;; TELEX: 200154
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 930
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
PCT-US93-05366-2

Query Match 25.4%; Score 127; DB 5; Length 930;
Best Local Similarity 100.0%; Pred. No. 2.5e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTTATTAGTCATCGCTATTACCATGGTGATCGGG 76
Db 469 GACTTTCCTACTTGGCAGTACATCTACGTTATTAGTCATCGCTATTACCATGGTGATCGGG 528

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTACCGGGATTTCCTCAAGTCTC 136
Db 529 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTACCGGGATTTCCTCAAGTCTC 588

QY 137 CACCCCA 143
Db 589 CACCCCA 595

RESULT 12
PCT-US93-05366-4
; Sequence 4, Application PC/TUS9305366
; GENERAL INFORMATION:
; APPLICANT: Robert J. Debs
; APPLICANT: Ning Zhu
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR IN VIVO
; TITLE OF INVENTION: GENE THERAPY
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05366
; FILING DATE: 04-JUN-1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/992,687
; FILING DATE: December 17, 1992
; APPLICATION NUMBER: 07/927,200
; FILING DATE: August 6, 1992
; APPLICATION NUMBER: 07/894,498
; FILING DATE: June 4, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Barbara Rae-Venter
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: 05935/005US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 854-5277
; TELEFAX: (415) 854-0875

TELEX: 200154
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 930
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US93-05386-4

Query Match 25.4%; Score 127; DB 5; Length 930;
Best Local Similarity 100.0%; Pred. No. 2.5e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTGCGG 76
DB 469 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTGCGG 528
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTTCCAAGTCTC 136
DB 529 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTTCCAAGTCTC 588
QY 137 CACCCCA 143
DB 589 CACCCCA 595

RESULT 13
US-09-310-842-1
Sequence 1, Application US/09310842A
Patent No. 6451593
GENERAL INFORMATION:
APPLICANT: Wittig, Prof. Burghardt
APPLICANT: Jungmans, Claas
TITLE OF INVENTION: Design Principle for Constructing Expression Constructs for Gene
TITLE OF INVENTION: Therapy
FILE REFERENCE: XI 597/99
CURRENT APPLICATION NUMBER: US/09/310,842A
CURRENT FILING DATE: 1999-05-12
EARLIER APPLICATION NUMBER: DE 196 48 625.4
EARLIER FILING DATE: 1996-11-13
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 1078
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: gene (1078)
LOCATION: (1)..(1078)
OTHER INFORMATION: bcr3=abl2; Oligo DNA Dumbbell
FEATURE:
NAME/KEY: misc binding
LOCATION: (1)..(2)
OTHER INFORMATION: intramolecular binding site; the T-nucleotides at position 1 to
OTHER INFORMATION: 2 can be modified with amino or caroxy features
FEATURE:
NAME/KEY: misc binding
LOCATION: (1077)..(1078)
OTHER INFORMATION: intramolecular binding site; the T-nucleotides at position 1077
OTHER INFORMATION: to 1078 can be modified with amino or caroxy features
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Strandedness: both; nucleic
OTHER INFORMATION: acid (linear), Hypothetical: No. 6451593 anti1-sense: No
US-09-310-842-1

Query Match 25.4%; Score 127; DB 4; Length 1078;
Best Local Similarity 100.0%; Pred. No. 2.5e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTGCGG 76
DB 352 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTGCGG 411

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTTCCAAGTCTC 136
DB 412 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTTCCAAGTCTC 471
QY 137 CACCCCA 143
DB 472 CACCCCA 478

RESULT 14
US-08-522-841-7
Sequence 7, Application US/08522841
Patent No. 5919652
GENERAL INFORMATION:
APPLICANT: Pang, Shen
APPLICANT: Belledgrun, Arie S.
TITLE OF INVENTION: No. 5919652el Nucleic Acid Molecules Comprising
TITLE OF INVENTION: the Prostate Specific Antigen (PSA) Promoter and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant and Gould
STREET: 11150 Santa Monica Blvd., Suite 400
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/522,841
FILING DATE: 01-SEP-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Adriano, Sarah B.
REGISTRATION NUMBER: 34,470
REFERENCE/DOCKET NUMBER: 30435.26US11
TELECOMMUNICATION INFORMATION:
TELEPHONE: 310-445-1140
TELEFAX: 310-445-9031
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1215 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-08-522-841-7

Query Match 25.4%; Score 127; DB 2; Length 1215;
Best Local Similarity 100.0%; Pred. No. 2.4e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTGCGG 76
DB 333 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTGCGG 392
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTTCCAAGTCTC 136
DB 393 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTTCCAAGTCTC 452
QY 137 CACCCCA 143
DB 453 CACCCCA 459

RESULT 15
US-08-522-841-3
Sequence 3, Application US/08522841.
Patent No. 5919652

GENERAL INFORMATION:
APPLICANT: Pang, Shen
APPLICANT: Beldegrun, Arie S.
TITLE OF INVENTION: No. 5919652el Nucleic Acid Molecules Comprising
TITLE OF INVENTION: the Prostate Specific Antigen (PSA) Promoter and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant and Gould
STREET: 11150 Santa Monica Blvd., Suite 400
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/522,841
FILING DATE: 01-SEP-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Adriano, Sarah B.
REGISTRATION NUMBER: 34,470
REFERENCE/DOCKET NUMBER: 30435.26US11
TELECOMMUNICATION INFORMATION:
TELEPHONE: 310-445-1140
TELEFAX: 310-445-9031
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1216 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-522-841-3

Query Match 25.4%; Score 127; DB 2; Length 1216;
Best Local Similarity 100.0%; Pred. NO. 2.4e-57;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 GACTTTCTTCTTGGCAGTACATCTCTAGTATTAGTCATCGCTATTACCATGGTGATGCGG 76
Db 333 GACTTTCTTCTTGGCAGTACATCTCTAGTATTAGTCATCGCTATTACCATGGTGATGCGG 392
QY 77 TTTTGGCAGTACATCAATGGCGGTGATAGCGGTTTGACTCAGCGGGATTTCCTCAAGTCTC 136
Db 393 TTTTGGCAGTACATCAATGGCGGTGATAGCGGTTTGACTCAGCGGGATTTCCTCAAGTCTC 452
QY 137 CACCCCA 143
Db 453 CACCCCA 459

Search completed: February 15, 2004, 19:10:12
Job time : 40.7706 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 16:21:35 ; Search time 210.032 Seconds
(without alignments)
8769.219 Million cell updates/sec

Title: US-09-921-143-36_COPY_1_500

Perfect score: 500
Sequence: 1 aagcttgacattatgcagt.....cacctgggtggatgc 500

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 2449703 seqs, 1841816367 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	500	100.0	5283	13	US-09-921-143-36
2	127	25.4	351	13	US-10-268-927-35
3	127	25.4	412	11	US-09-911-020A-1
4	127	25.4	544	13	US-10-268-927-33
5	127	25.4	600	9	US-09-895-828-425
6	127	25.4	600	15	US-10-114-666-425
7	127	25.4	619	10	US-09-965-703-68
8	127	25.4	645	13	US-10-437-830-3
9	127	25.4	851	12	US-10-465-156-1
10	127	25.4	884	15	US-10-152-363A-69
11	127	25.4	930	15	US-10-206-747-2
12	127	25.4	930	15	US-10-206-747-4
13	127	25.4	987	13	US-09-845-416-33
14	127	25.4	996	8	US-08-900-220C-25
15	127	25.4	996	10	US-09-151-999-25

16	127	25.4	996	11	US-09-187-387-25	Sequence 25, Appl
17	127	25.4	996	11	US-09-827-110-25	Sequence 25, Appl
18	127	25.4	996	13	US-09-238-243-25	Sequence 25, Appl
19	127	25.4	1070	15	US-10-205-534-6	Sequence 6, Appl
20	127	25.4	1078	15	US-10-228-811-1	Sequence 1, Appl
21	127	25.4	1104	10	US-09-845-899A-6	Sequence 6, Appl
22	127	25.4	1318	15	US-10-228-811-3	Sequence 3, Appl
23	127	25.4	1417	15	US-10-228-811-5	Sequence 5, Appl
24	127	25.4	1567	12	US-10-431-121A-3	Sequence 3, Appl
25	127	25.4	1645	15	US-10-228-811-2	Sequence 2, Appl
26	127	25.4	1715	9	US-09-886-942-10	Sequence 10, Appl
27	127	25.4	1767	9	US-09-886-942-8	Sequence 8, Appl
28	127	25.4	1767	9	US-09-886-942-19	Sequence 19, Appl
29	127	25.4	1787	12	US-10-364-049-900	Sequence 900, App
30	127	25.4	1813	9	US-09-794-975-8	Sequence 8, Appl
31	127	25.4	1870	15	US-10-228-811-4	Sequence 4, Appl
32	127	25.4	2335	15	US-10-078-167-2	Sequence 2, Appl
33	127	25.4	3360	15	US-10-023-969-4	Sequence 4, Appl
34	127	25.4	3426	13	US-10-136-837-2	Sequence 2, Appl
35	127	25.4	3487	15	US-10-303-683-16	Sequence 16, Appl
36	127	25.4	3547	12	US-10-369-121-43	Sequence 43, Appl
37	127	25.4	3589	13	US-10-136-837-4	Sequence 4, Appl
38	127	25.4	3589	15	US-10-234-406-1	Sequence 1, Appl
39	127	25.4	3609	15	US-10-234-406-3	Sequence 3, Appl
40	127	25.4	3610	11	US-09-194-949-1	Sequence 1, Appl
41	127	25.4	3788	9	US-09-392-462-1	Sequence 1, Appl
42	127	25.4	3788	9	US-09-540-991-1	Sequence 1, Appl
43	127	25.4	3788	13	US-10-291-041-1	Sequence 1, Appl
44	127	25.4	3796	13	US-10-241-332-32	Sequence 32, Appl
45	127	25.4	3796	14	US-10-127-391-32	Sequence 32, Appl

ALIGNMENTS

RESULT 1
US-09-921-143-36
; Sequence 36, Application US/09921143
; Publication No. US20030215921A1
; GENERAL INFORMATION:
; APPLICANT: Coleman, Timothy
; TITLE OF INVENTION: Vascular Endothelial Growth Factor-2
; FILE REFERENCE: PFI12P6
; CURRENT APPLICATION NUMBER: US/09/921,143
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/223,276
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent version 3.0
; SEQ ID NO 36
; LENGTH: 5283
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-921-143-36

Query Match 100.0%; Score 500; DB 13; Length 5283;
Best Local Similarity 100.0%; Pred. No. 8.8e-272;
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	AAGCTTGACCTTATGGGACTTTCTACTTGGCAGTACATCTAGTATTAGTATCGCTAT	60
Db	1	AAGCTTGACCTTATGGGACTTTCTACTTGGCAGTACATCTAGTATTAGTATCGCTAT	60
Qy	61	TACATGGTGATCGGTTTGGCAGTACATCAATGGCGTGGATACGCGTTTGACATCACG	120
Db	61	TACATGGTGATCGGTTTGGCAGTACATCAATGGCGTGGATACGCGTTTGACATCACG	120
Qy	121	GGGATTTCGAAGTCTCCACCCCACTGACGCTCAATGGGAGTTGTTTGGACCAAAATCA	180
Db	121	GGGATTTCGAAGTCTCCACCCCACTGACGCTCAATGGGAGTTGTTTGGACCAAAATCA	180
Qy	181	ACGAGATTTCGAAATGTCGTAACACTCCGCCCACTTCACGCAATGGCGGTAGGCG	240

Db 181 ACAGACTTCCAAAATGTCGTAAACAACACTCGCCGCCCAATTGACGCAAAATGGCGGTAGGCG 240
Qy 241 PACATGCTTATGTACGCTGAGTTAGCAACATGCTTATAAGGAGAGAAAAGCACCGTG 300
Db 241 AACATGCTTATGTACGCTGAGTTAGCAACATGCTTATAAGGAGAGAAAAGCACCGTG 300
Qy 301 CATGCCGATTGGTGGAGTAAGGTGCTATGATTCGTGGTATGATCGTGCCTTTGTAGGAAG 360
Db 301 CATGCCGATTGGTGGAGTAAGGTGCTATGATTCGTGGTATGATCGTGCCTTTGTAGGAAG 360
Qy 361 GCAACAGACGGGTCTAAACACGATTGGACGACCACTCAATTCGCGATTGACAGATATT 420
Db 361 GCAACAGACGGGTCTAAACACGATTGGACGACCACTCAATTCGCGATTGACAGATATT 420
Qy 421 GTATTTAAGTGGCCAGCTCGATACAAATAAGCCCAATTTGACCAATTCACCAATTGGTGTG 480
Db 421 GTATTTAAGTGGCCAGCTCGATACAAATAAGCCCAATTTGACCAATTCACCAATTGGTGTG 480
Qy 481 CACCTGGGTGGATCGATC 500
Db 481 CACCTGGGTGGATCGATC 500

RESULT 2

US-10-268-927-35

; Sequence 35, Application US/10268927

; Publication No. US20030175693A1

; GENERAL INFORMATION:

; APPLICANT: Main-Hobson, Simon

; APPLICANT: Blancou, Philippe

; APPLICANT: Chenciner, Nicole

; TITLE OF INVENTION: HIV RECOMBINANT VACCINE

; FILE REFERENCE: 03495.0239-00000

; CURRENT APPLICATION NUMBER: US/10/268,927

; CURRENT FILING DATE: 2002-10-11

; PRIOR APPLICATION NUMBER: 60/328,449

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 35

; LENGTH: 351

; TYPE: DNA

; ORGANISM: Artificial

; FEATURE:

; OTHER INFORMATION: recombinant promoter

US-10-268-927-35

Query Match 25.4%; Score 127; DB 13; Length 351;
Best Local Similarity 100.0%; Pred. No. 5.1e-61;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTATGCGG 76
Db 74 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTATGCGG 133
Qy 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 136
Db 134 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 193
Qy 137 CACCCCA 143
Db 194 CACCCCA 200

RESULT 3

US-09-911-020A-1

; Sequence 1, Application US/09911020A

; Publication No. US20030040100A1

; GENERAL INFORMATION:

; APPLICANT: BROUGH, DOUGLAS E

; APPLICANT: KOVESDI, IMRE

; TITLE OF INVENTION: CELL FOR THE PROPAGATION OF ADENOVIRAL VECTORS

; FILE REFERENCE: 209459

; CURRENT APPLICATION NUMBER: US/09/911,020A
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 412
; TYPE: DNA
; ORGANISM: Cytomegalovirus
; US-09-911-020A-1

Query Match 25.4%; Score 127; DB 11; Length 412;
Best Local Similarity 100.0%; Pred. No. 5.1e-61;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTATGCGG 76
Db 256 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTATGCGG 315
Qy 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 136
Db 316 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 375
Qy 137 CACCCCA 143
Db 376 CACCCCA 382

RESULT 4

US-10-268-927-33

; Sequence 33, Application US/10268927

; Publication No. US20030175693A1

; GENERAL INFORMATION:

; APPLICANT: Main-Hobson, Simon

; APPLICANT: Blancou, Philippe

; APPLICANT: Chenciner, Nicole

; TITLE OF INVENTION: HIV RECOMBINANT VACCINE

; FILE REFERENCE: 03495.0239-00000

; CURRENT APPLICATION NUMBER: US/10/268,927

; CURRENT FILING DATE: 2002-10-11

; PRIOR APPLICATION NUMBER: 60/328,449

; PRIOR FILING DATE: 2001-10-12

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 33

; LENGTH: 544

; TYPE: DNA

; ORGANISM: Artificial

; FEATURE:

; OTHER INFORMATION: recombinant promoter

US-10-268-927-33

Query Match 25.4%; Score 127; DB 13; Length 544;
Best Local Similarity 100.0%; Pred. No. 5.1e-61;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTATGCGG 76
Db 267 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGTATGCGG 326
Qy 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 136
Db 327 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 386
Qy 137 CACCCCA 143
Db 387 CACCCCA 393

RESULT 5

US-09-895-828-425

; Sequence 425, Application US/09895828

; Patent No. US2002009012A1

; GENERAL INFORMATION:

APPLICANT: Wang, Tongtong
APPLICANT: McNeill, Patricia D.
APPLICANT: Watanabe, Yoshihiro
APPLICANT: Carter, Darrick
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.539
CURRENT APPLICATION NUMBER: US/09/895,828
CURRENT FILING DATE: 2001-06-28
NUMBER OF SEQ ID NOS: 473
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 425
LENGTH: 600
TYPE: DNA
ORGANISM: Homo sapiens
US-09-895-828-425

Query Match 25.4%; Score 127; DB 9; Length 600;
Best Local Similarity 100.0%; Pred. No. 5.1e-61;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGGG 76
Db 278 GACTTTCCTACTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGGG 337

QY 77 TTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 136
Db 338 TTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 397

QY 137 CACCCCA 143
Db 398 CACCCCA 404

RESULT 6
US-10-114-666-425
Sequence 425, Application US/10114666
Publication No. US20030103994A1
GENERAL INFORMATION:
APPLICANT: Watanabe, Yoshihiro
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.539C1
CURRENT APPLICATION NUMBER: US/10/114,666
CURRENT FILING DATE: 2002-04-01
NUMBER OF SEQ ID NOS: 479
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 425
LENGTH: 600
TYPE: DNA
ORGANISM: Homo sapiens
US-10-114-666-425

Query Match 25.4%; Score 127; DB 15; Length 600;
Best Local Similarity 100.0%; Pred. No. 5.1e-61;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGGG 76
Db 278 GACTTTCCTACTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGGG 337

QY 77 TTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 136
Db 338 TTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 397

QY 137 CACCCCA 143
Db 398 CACCCCA 404

RESULT 7
US-09-965-703-68
Sequence 68, Application US/09965703
Patent No. US20020119521A1
GENERAL INFORMATION:
APPLICANT: Rohm and Haas Company
APPLICANT: Palli, Subba Reddy
APPLICANT: Kapitskaya, Marianna zinovjevna
APPLICANT: Cress, Dean Ervin
TITLE OF INVENTION: No. US20020119521A1el Bodysone Receptor-Based Inducible Gene E
FILE REFERENCE: A01020B
CURRENT APPLICATION NUMBER: US/09/965,703
CURRENT FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/191,355
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/269,799
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: PCT/US01/09050
PRIOR FILING DATE: 2001-03-21
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn version 3.1
SEQ ID NO 68
LENGTH: 619
TYPE: DNA
ORGANISM: Cytomegalovirus
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: No. US20020119521A1el Sequence
US-09-965-703-68

Query Match 25.4%; Score 127; DB 10; Length 619;
Best Local Similarity 100.0%; Pred. No. 5.1e-61;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGGG 76
Db 254 GACTTTCCTACTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATCGGG 313

QY 77 TTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 136
Db 314 TTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAAGTCTC 373

QY 137 CACCCCA 143
Db 374 CACCCCA 380

RESULT 8
US-10-437-830-3
Sequence 3, Application US/10437830
Publication No. US20030180949A1
GENERAL INFORMATION:
APPLICANT: JOHN LEVY
TITLE OF INVENTION: AVOIDANCE OF UNDESIRABLE REPLICATION
TITLE OF INVENTION: INTERMEDIATES IN PLASMID PROPAGATION
FILE REFERENCE: CT1MM.014DV1
CURRENT APPLICATION NUMBER: US/10/437,830
CURRENT FILING DATE: 2003-05-13
PRIOR APPLICATION NUMBER: 09/715,835
PRIOR FILING DATE: 2000-11-16
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 645
TYPE: DNA
ORGANISM: CYTOMEGALOVIRUS
US-10-437-830-3

Query Match 25.4%; Score 127; DB 13; Length 645;
Best Local Similarity 100.0%; Pred. No. 5.1e-61;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTATCGTATTCAGTATTCACCATGCTGATCGGG 76
DB 431 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTATTCAGTATTCACCATGCTGATCGGG 490
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGATTTCCAAAGTCTC 136
DB 491 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGATTTCCAAAGTCTC 550
QY 137 CACCCCA 143
DB 551 CACCCCA 557

RESULT 9
US-10-465-156-1
; Sequence 1, Application US/10465156
; Publication No. US20030232414A1
; GENERAL INFORMATION:
; APPLICANT: Moore, Margaret Dow
; TITLE OF INVENTION: HYBRID VECTOR HAVING A CYTOMEGALOVIRUS
; TITLE OF INVENTION: ENHANCER AND MYELOPROLIFERATIVE SARCOMA VIRUS PROMOTER
; FILE REFERENCE: 02-15
; CURRENT APPLICATION NUMBER: US/10/465,156
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: 60/389,612
; PRIOR FILING DATE: 2002-06-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 851
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: hybrid cytomegalovirus and myeloproliferative
; OTHER INFORMATION: sarcoma virus regulatory sequence
US-10-465-156-1

Query Match 25.4%; Score 127; DB 12; Length 851;
Best Local Similarity 100.0%; Pred. No. 5.2e-61;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTATTCAGTATTCACCATGCTGATCGGG 76
DB 223 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTATTCAGTATTCACCATGCTGATCGGG 282
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGATTTCCAAAGTCTC 136
DB 283 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGATTTCCAAAGTCTC 342
QY 137 CACCCCA 143
DB 343 CACCCCA 349

RESULT 10
US-10-152-363A-69
; Sequence 69, Application US/10152363A
; Publication No. US20030103986A1
; GENERAL INFORMATION:
; APPLICANT: Rixon, Mark W.
; APPLICANT: Gross, Jane A.
; TITLE OF INVENTION: TACI-Immunoglobulin Fusion Proteins
; FILE REFERENCE: 01-20
; CURRENT APPLICATION NUMBER: US/10/152,363A
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 60/293,343
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 69
; LENGTH: 884
; TYPE: DNA
; ORGANISM: Artificial Sequence

; FEATURE:
; OTHER INFORMATION: CMV enhancer/MPSV LTR promoter construct.
US-10-152-363A-69
Query Match 25.4%; Score 127; DB 15; Length 884;
Best Local Similarity 100.0%; Pred. No. 5.2e-61;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTATTCAGTATTCACCATGCTGATCGGG 76
DB 256 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTATTCACCATGCTGATCGGG 315
QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGATTTCCAAAGTCTC 136
DB 316 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGATTTCCAAAGTCTC 375
QY 137 CACCCCA 143
DB 376 CACCCCA 382
RESULT 11
US-10-206-747-2
; Sequence 2, Application US/10206747
; Publication No. US20030109475A1
; GENERAL INFORMATION:
; APPLICANT: (i) APPLICANT: Zhu, Ning
; APPLICANT: (i) APPLICANT: Zhu, Ning
; APPLICANT: Zhu, Ning
; TITLE OF INVENTION: Methods and Compositions for In Vivo
; Gene Therapy
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/206,747
; FILING DATE: 11-Feb-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/809,292
; FILING DATE: 17-DEC-1991
; APPLICATION NUMBER: US 07/894,498
; FILING DATE: 04-JUN-1992
; APPLICATION NUMBER: US 07/927,200
; FILING DATE: 06-AUG-1992
; APPLICATION NUMBER: US 07/972,135
; FILING DATE: 05-NOV-1992
; APPLICATION NUMBER: US 07/992,687
; FILING DATE: 17-DEC-1992
; APPLICATION NUMBER: WO PCT/US92/11004
; FILING DATE: 17-DEC-1992
; APPLICATION NUMBER: WO PCT/US93/05366
; FILING DATE: 04-JUN-1993
; APPLICATION NUMBER: US 08/464,899
; FILING DATE: 22-JAN-1996
; APPLICATION NUMBER: US 09/132,391
; FILING DATE: 08-NOV-2003
; ATTORNEY/AGENT INFORMATION:
; NAME: Mycroft, Frank J.
; REGISTRATION NUMBER: 46,946
; REFERENCE/DOCKET NUMBER: 023070-064833US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200

```
;
; TELEFAX: (415) 576-0300
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 930 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA
;   SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-206-747-2

Query Match      25.4%; Score 127; DB 15; Length 930;
Best Local Similarity 100.0%; Pred. No. 5.2e-61;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGG 76
Db 469 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGG 528

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTCCAAAGTCTC 136
Db 529 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTCCAAAGTCTC 588

QY 137 CACCCCA 143
Db 589 CACCCCA 595

RESULT 12
US-10-206-747-4
; Sequence 4, Application US/10206747
; Publication No. US20030109475A1
; GENERAL INFORMATION:
; APPLICANT: (i) APPLICANT: Zhu, Ning
;              (ii) APPLICANT: Zhu, Ning
;              Zhu, Ning
; TITLE OF INVENTION: Methods and Compositions for In Vivo
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/206,747
; FILING DATE: 11-Feb-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/809,292
; FILING DATE: 17-DEC-1991
; APPLICATION NUMBER: US 07/894,498
; FILING DATE: 04-JUN-1992
; APPLICATION NUMBER: US 07/927,200
; FILING DATE: 06-AUG-1992
; APPLICATION NUMBER: US 07/972,135
; FILING DATE: 05-NOV-1992
; APPLICATION NUMBER: US 07/992,687
; FILING DATE: 17-DEC-1992
; APPLICATION NUMBER: WO PCT/US92/11004
; FILING DATE: 17-DEC-1992
; APPLICATION NUMBER: WO PCT/US93/05366
; FILING DATE: 04-JUN-1993
; APPLICATION NUMBER: US 08/464,899
; FILING DATE: 22-JAN-1996
```

```
;
; APPLICATION NUMBER: US 09/132,391
; FILING DATE: 08-No. US20030109475A1-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Mycroft, Frank J.
; REGISTRATION NUMBER: 46,946
; REFERENCE/DOCKET NUMBER: 023070-064833US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 930 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA
;   SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-206-747-4

Query Match      25.4%; Score 127; DB 15; Length 930;
Best Local Similarity 100.0%; Pred. No. 5.2e-61;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGG 76
Db 469 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGG 528

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTCCAAAGTCTC 136
Db 529 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTCCAAAGTCTC 588

QY 137 CACCCCA 143
Db 589 CACCCCA 595

RESULT 13
US-09-845-416-33
; Sequence 33, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DEL142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 987
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-845-416-33

Query Match      25.4%; Score 127; DB 13; Length 987;
Best Local Similarity 100.0%; Pred. No. 5.2e-61;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGG 76
Db 417 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGG 476

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTCCAAAGTCTC 136
Db 477 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCAGCGGGATTCCAAAGTCTC 536

QY 137 CACCCCA 143
Db 537 CACCCCA 543
```

RESULT 14

US-08-900-220C-25
; Sequence 25, Application US/08900220C
; Publication No. US20020045206A1

GENERAL INFORMATION:

APPLICANT: Miao, Ningning
; Wang, Monica
; Mahanthappa, Nagesh K.
; Pang, Kevin
; Jin, Ping

TITLE OF INVENTION: Method of Treating Dopaminergic and
GABA-nergic Disorders

NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: POLEY, HOAG & ELIOT LLP
STREET: ONE POST OFFICE SQUARE
CITY: Boston
STATE: MA

COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII (text)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/900,220C
FILING DATE: 24-Jul-1997
CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: ONV-044.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 996 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: nucleic acid

SEQUENCE DESCRIPTION: SEQ ID NO: 25:

US-08-900-220C-25

Query Match 25.4%; Score 127; DB 8; Length 996;

Best Local Similarity 100.0%; Pred. No. 5.2e-61;

Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 GACITTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTATCGCG 76

Db 593 GACITTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTATCGCG 652

Qy 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGTTTGACTCACGGGGATTTCCAAAGTCTC 136

Db 653 TTTTGGCAGTACATCAATGGCGGTGGATAGCGTTTGACTCACGGGGATTTCCAAAGTCTC 712

Qy 137 CACCCCA 143

Db 713 CACCCCA 719

RESULT 15

US-09-151-999-25

; Sequence 25, Application US/09151999

; Patent No. US20020151460A1

GENERAL INFORMATION:

APPLICANT: Wang, Elizabeth

; TITLE OF INVENTION: REGULATION OF EPITHELIAL TISSUE BY HEDGEHOG-LIKE

; TITLE OF INVENTION: POLYPEPTIDES, AND FORMULATIONS AND USES RELATED THERETO

; FILE REFERENCE: ONV-031.02
; CURRENT APPLICATION NUMBER: US/09/151,999
; CURRENT FILING DATE: 1998-08-11

; EARLIER APPLICATION NUMBER: 08/955,552

; EARLIER FILING DATE: 1997-10-20

; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 25

; LENGTH: 996

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: gene

; OTHER INFORMATION: activation construct

US-09-151-999-25

Query Match 25.4%; Score 127; DB 10; Length 996;

Best Local Similarity 100.0%; Pred. No. 5.2e-61;

Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 GACITTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTATCGCG 76

Db 593 GACITTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTATCGCG 652

Qy 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGTTTGACTCACGGGGATTTCCAAAGTCTC 136

Db 653 TTTTGGCAGTACATCAATGGCGGTGGATAGCGTTTGACTCACGGGGATTTCCAAAGTCTC 712

Qy 137 CACCCCA 143

Db 713 CACCCCA 719

Search completed: February 15, 2004, 22:26:45

Job time : 213.032 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 21:50:21 ; Search time 1316.95 Seconds
(without alignments)
9227.587 Million cell updates/sec

Title: US-09-921-143-36_COPY_1_500

Perfect score: 500

Sequence: 1 aagcttgacattatgcgact.....cacctgggtggatcgatc 500

Scoring table: OLIGO NUC

Gapop_60.0, Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size : 0

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

```

1: em_estba:*
2: em_estum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_eston:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_rtr:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	127	25.4	187	12	BI489763
2	127	25.4	191	12	BI489620
3	127	25.4	193	12	BI490268
4	127	25.4	193	12	BI490463

C	5	127	25.4	480	29	B2996651	B2996651
C	6	127	25.4	500	29	B2680401	B2680401
C	7	127	25.4	610	28	B68191	B68191
C	8	127	25.4	619	9	AV645385	AV645385
C	9	127	25.4	634	9	AV645390	AV645390
C	10	127	25.4	637	9	AV645392	AV645392
C	11	127	25.4	640	9	AV645393	AV645393
C	12	127	25.4	648	9	AV681492	AV681492
C	13	127	25.4	655	9	AV681500	AV681500
C	14	127	25.4	657	9	AV681507	AV681507
C	15	127	25.4	661	9	AV681509	AV681509
C	16	127	25.4	663	9	AV681529	AV681529
C	17	127	25.4	681	9	AV681512	AV681512
C	18	127	25.4	681	9	AV681513	AV681513
C	19	127	25.4	684	9	AV681526	AV681526
C	20	127	25.4	694	9	AV645326	AV645326
C	21	127	25.4	733	9	AV681507	AV681507
C	22	127	25.4	744	9	AV681465	AV681465
C	23	127	25.4	764	29	B2817918	B2817918
C	24	127	25.4	768	9	AV681525	AV681525
C	25	127	25.4	772	9	AV681495	AV681495
C	26	127	25.4	776	9	AV681505	AV681505
C	27	127	25.4	783	9	AV681502	AV681502
C	28	127	25.4	786	9	AV681504	AV681504
C	29	127	25.4	802	9	AV681486	AV681486
C	30	127	25.4	803	9	AV681483	AV681483
C	31	127	25.4	806	9	AV681464	AV681464
C	32	127	25.4	807	9	AV681462	AV681462
C	33	127	25.4	810	9	AV681487	AV681487
C	34	127	25.4	811	9	AV681506	AV681506
C	35	127	25.4	816	9	AV681523	AV681523
C	36	127	25.4	821	9	AV681519	AV681519
C	37	127	25.4	834	29	B2667422	B2667422
C	38	127	25.4	888	13	BX330112	BX330112
C	39	127	25.4	905	13	BX459694	BX459694
C	40	127	25.4	919	13	BX372036	BX372036
C	41	127	25.4	920	13	BX393430	BX393430
C	42	127	25.4	932	13	BX393879	BX393879
C	43	127	25.4	948	13	BX364612	BX364612
C	44	127	25.4	956	13	BX359183	BX359183
C	45	127	25.4	963	13	BX342541	BX342541

ALIGNMENTS

BI489763 187 bp mRNA linear EST 28-AUG-2001
603032280F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5173506 5',
mRNA sequence.
BI489763
EST.
BI489763.1 GI:15328991
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 187)
NIH-MGC http://mgc.ncbi.nlm.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaabs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing By: The I.M.A.G.E. Consortium (LLNL)
Clone Distribution: Incyte Genomics, Inc.
Found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL1431 row: p column: 19
High quality sequence stop: 187.

```

FEATURES
source
Location/Qualifiers
1..187
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5173506"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/note="Organ: pooled brain, lung, testis; Vector:
pcmv-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."
BASE COUNT 38 a 44 c 48 g 57 t
ORIGIN

Query Match 25.4%; Score 127; DB 12; Length 187;
Best Local Similarity 100.0%; Pred. No. 5.5e-58;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTCCCTACTGCGAGTACATCTACGTATTAGTATTCATCGCTATTACCATGCGTGGCGG 76
Db 39 GACTTCCCTACTGCGAGTACATCTACGTATTAGTATTCATCGCTATTACCATGCGTGGCGG 98

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATGAGCGGTTTGACTCACGGGGATTTCACAGTCTC 136
Db 99 TTTTGGCAGTACATCAATGGCGGTGGATGAGCGGTTTGACTCACGGGGATTTCACAGTCTC 158

QY 137 CACCCCA 143
Db 159 CACCCCA 165

RESULT 2
BI489620
LOCUS 603032080F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5173483 5',
DEFINITION mRNA sequence.
ACCESSION BI489620
VERSION BI489620.1 GI:15328848
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 191)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Prepared by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11431 row: 0 column: 20
High quality sequence stop: 191.
Location/Qualifiers
1..191
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5173483"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/note="Organ: pooled brain, lung, testis; Vector:
pcmv-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."
BASE COUNT 39 a 44 c 50 g 58 t
ORIGIN

Query Match 25.4%; Score 127; DB 12; Length 191;
Best Local Similarity 100.0%; Pred. No. 5.5e-58;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTCCCTACTGCGAGTACATCTACGTATTAGTATTCATCGCTATTACCATGCGTGGCGG 76
Db 38 GACTTCCCTACTGCGAGTACATCTACGTATTAGTATTCATCGCTATTACCATGCGTGGCGG 97

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATGAGCGGTTTGACTCACGGGGATTTCACAGTCTC 136
Db 98 TTTTGGCAGTACATCAATGGCGGTGGATGAGCGGTTTGACTCACGGGGATTTCACAGTCTC 157

QY 137 CACCCCA 143
Db 158 CACCCCA 164

RESULT 3
BI490268
LOCUS 603032080R1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5173483 3',
DEFINITION mRNA sequence.
ACCESSION BI490268
VERSION BI490268.1 GI:15329483
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 193)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Prepared by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11431 row: 0 column: 20
High quality sequence stop: 193.
Location/Qualifiers
1..193
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5173483"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/note="Organ: pooled brain, lung, testis; Vector:
pcmv-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.

```

Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

BASE COUNT 59 a 50 c 46 g 38 t

ORIGIN

Query Match 25.4%; Score 127; DB 12; Length 193;
Best Local Similarity 100.0%; Pred. No. 5.6e-58;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTTATAGTACGCTATACCATGGTATGGCG 76
Db 154 GACTTTCCTACTTGGCAGTACATCTACGTTATAGTACGCTATACCATGGTATGGCG 95

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGTTTGGTACCTACGGGGATTTCCAAAGTCTC 136
Db 94 TTTTGGCAGTACATCAATGGCGGTGGATAGCGTTTGGTACCTACGGGGATTTCCAAAGTCTC 35

QY 137 CACCCCA 143
Db 34 CACCCCA 28

RESULT 4
BI490463/c

LOCUS 603032280R1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5173506 3',
DEFINITION mRNA sequence.

ACCESSION BI490463

VERSION BI490463.1 GI:15329691

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 193)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-f@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M11431 row: p column: 19
High quality sequence stop: 193.

FEATURES
Location/Qualifiers
1..193
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5173506"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORE6; Site_1: Not; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

BASE COUNT 59 a 50 c 46 g 38 t

ORIGIN

Query Match 25.4%; Score 127; DB 12; Length 193;
Best Local Similarity 100.0%; Pred. No. 5.6e-58;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTTATAGTACGCTATACCATGGTATGGCG 76
Db 154 GACTTTCCTACTTGGCAGTACATCTACGTTATAGTACGCTATACCATGGTATGGCG 95

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGTTTGGTACCTACGGGGATTTCCAAAGTCTC 136
Db 94 TTTTGGCAGTACATCAATGGCGGTGGATAGCGTTTGGTACCTACGGGGATTTCCAAAGTCTC 35

QY 137 CACCCCA 143
Db 34 CACCCCA 28

RESULT 5
BZ996651/c

LOCUS PUBKH77TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa074N10,
DEFINITION genomic survey sequence.

ACCESSION BZ996651

VERSION BZ996651.1 GI:29240068

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

REFERENCE 1 (bases 1 to 480)
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
TITLE Whitealaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick
A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennetzen,J.
JOURNAL Maize Genomics Consortium
COMMENT Unpublished
Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.

FEATURES
Location/Qualifiers
1..480
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTa074N10"
/clone_lib="ZM_0.6_1.0_KB"
/note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
Cor selected genomic DNA library"

BASE COUNT 133 a 108 c 118 g 121 t

ORIGIN

Query Match 25.4%; Score 127; DB 29; Length 480;
Best Local Similarity 100.0%; Pred. No. 6.5e-58;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTTATAGTACGCTATACCATGGTATGGCG 76
Db 164 GACTTTCCTACTTGGCAGTACATCTACGTTATAGTACGCTATACCATGGTATGGCG 105

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGTTTGGTACCTACGGGGATTTCCAAAGTCTC 136
Db 104 TTTTGGCAGTACATCAATGGCGGTGGATAGCGTTTGGTACCTACGGGGATTTCCAAAGTCTC 45

QY 137 CACCCCA 143
Db 44 CACCCCA 38

RESULT 6
BZ680401

LOCUS 500 bp DNA linear GSS 05-FEB-2003

Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: T7
Class: BAC ends.

FEATURES
Location/Qualifiers
1. 610
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="CDS:726476"
/db_xref="taxon:9606"
/clones="A-492C12"
/sex="Female"
/cell_type="Fibroblast"
/clone_lib="CIR78SK"
/note="Vector: pBAC1081; Site_1: HindIII; Site_2: HindIII;
CalTech Human BAC Library A"
146 a 159 c 165 g 140 t

BASE COUNT 146 a 159 c 165 g 140 t
ORIGIN

Query Match 25.4%; Score 127; DB 28; Length 610;
Best Local Similarity 100.0%; Pred. No. 6.8e-58;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTTCATCGCTATTACCATGTGTATCGGG 76
|||
Db 508 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTTCATCGCTATTACCATGTGTATCGGG 449
|||

QY 77 TTTTGGCAGTACATCAATGGCGGTGGATACGGTTTGACTTCACGGGGATTTCCAAAGTCTC 136
|||
Db 448 TTTTGGCAGTACATCAATGGCGGTGGATACGGTTTGACTTCACGGGGATTTCCAAAGTCTC 389
|||

QY 137 CACCCCA 143
|||
Db 388 CACCCCA 382
|||

RESULT 8
AV645385/c
LOCUS
DEFINITION GLA Homo sapiens cDNA clone GLAAAG02 3', mRNA sequence.
ACCESSION AV645385
VERSION AV645385.1 GI:9866399
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 619)
Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
21625106
11752456

REFERENCE
AUTHORS Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex. 45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn

TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

This clone is available at CHGC in Shanghai.

```

FEATURES
  source
    1..619
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clones="GLAAG02"
    /tissue_type="corresponding non cancerous liver tissue"
    /dev_stage="Adult"
    /lab_host="SOLR"
    /clone_lib="GLA"
    /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
    XhoI"

BASE COUNT  147 a  164 c  159 g  145 t      4 others
ORIGIN
Query Match      25.4%; Score 127; DB 9; Length 619;
Best Local Similarity 100.0%; Pred. No. 6.8e-58;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  17 GACTTTCCTACTTGGCAGTACATCATCGTATTAGTCATCGCTATTACCATGGTGATCGG 76
Db  385 GACTTTCCTACTTGGCAGTACATCATCGTATTAGTCATCGCTATTACCATGGTGATCGG 326

QY  77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGATTTCGAAGTCTC 136
Db  325 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGATTTCGAAGTCTC 266

QY  137 CACCCCA 143
Db  265 CACCCCA 259

RESULT 9
LOCUS      AV645390/c
DEFINITION AV645390 GLA Homo sapiens cDNA clone GLAAG07 3', mRNA sequence.
ACCESSION  AV645390
VERSION     AV645390.1 GI:9866404
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 634)
AUTHORS   Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
TITLE     Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
MEDLINE   21625106
PUBMED    11752456
COMMENT   Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
  source
    1..634
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clones="GLAAG07"
    /tissue_type="corresponding non cancerous liver tissue"
    /dev_stage="Adult"
    /lab_host="SOLR"
    /clone_lib="GLA"

BASE COUNT  147 a  164 c  159 g  145 t      4 others
ORIGIN
Query Match      25.4%; Score 127; DB 9; Length 619;
Best Local Similarity 100.0%; Pred. No. 6.8e-58;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  17 GACTTTCCTACTTGGCAGTACATCATCGTATTAGTCATCGCTATTACCATGGTGATCGG 76
Db  385 GACTTTCCTACTTGGCAGTACATCATCGTATTAGTCATCGCTATTACCATGGTGATCGG 326

QY  77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGATTTCGAAGTCTC 136
Db  325 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGATTTCGAAGTCTC 266

QY  137 CACCCCA 143
Db  265 CACCCCA 259

RESULT 9
LOCUS      AV645390/c
DEFINITION AV645390 GLA Homo sapiens cDNA clone GLAAG07 3', mRNA sequence.
ACCESSION  AV645390
VERSION     AV645390.1 GI:9866404
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 634)
AUTHORS   Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
TITLE     Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
MEDLINE   21625106
PUBMED    11752456
COMMENT   Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
  source
    1..637
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clones="GLAAG09"
    /tissue_type="corresponding non cancerous liver tissue"
    /dev_stage="Adult"
    /lab_host="SOLR"
    /clone_lib="GLA"
    /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
    XhoI"

BASE COUNT  157 a  157 c  154 g  157 t      12 others
ORIGIN
Query Match      25.4%; Score 127; DB 9; Length 637;
Best Local Similarity 100.0%; Pred. No. 6.8e-58;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  17 GACTTTCCTACTTGGCAGTACATCATCGTATTAGTCATCGCTATTACCATGGTGATCGG 76

```

```

/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT  152 a  158 c  160 g  156 t      8 others
ORIGIN
Query Match      25.4%; Score 127; DB 9; Length 634;
Best Local Similarity 100.0%; Pred. No. 6.8e-58;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  17 GACTTTCCTACTTGGCAGTACATCATCGTATTAGTCATCGCTATTACCATGGTGATCGG 76
Db  384 GACTTTCCTACTTGGCAGTACATCATCGTATTAGTCATCGCTATTACCATGGTGATCGG 325

QY  77 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGATTTCGAAGTCTC 136
Db  324 TTTTGGCAGTACATCAATGGCGGTGGATAGCGGTTTGACTCACGGGATTTCGAAGTCTC 265

QY  137 CACCCCA 143
Db  264 CACCCCA 258

RESULT 10
LOCUS      AV645392/c
DEFINITION AV645392 GLA Homo sapiens cDNA clone GLAAG09 3', mRNA sequence.
ACCESSION  AV645392
VERSION     AV645392.1 GI:9866406
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 637)
AUTHORS   Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
TITLE     Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
MEDLINE   21625106
PUBMED    11752456
COMMENT   Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
  source
    1..637
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clones="GLAAG09"
    /tissue_type="corresponding non cancerous liver tissue"
    /dev_stage="Adult"
    /lab_host="SOLR"
    /clone_lib="GLA"
    /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
    XhoI"

BASE COUNT  157 a  157 c  154 g  157 t      12 others
ORIGIN
Query Match      25.4%; Score 127; DB 9; Length 637;
Best Local Similarity 100.0%; Pred. No. 6.8e-58;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  17 GACTTTCCTACTTGGCAGTACATCATCGTATTAGTCATCGCTATTACCATGGTGATCGG 76

```

```

Db      385 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCAGTGTGATCGG 326
Qy      77 TTTTGGCAGTACATCAATGGCGGTGATAGCGTTTGACTCAGCGGGATTTCCAAAGTC 136
Db      325 TTTTGGCAGTACATCAATGGCGGTGATAGCGTTTGACTCAGCGGGATTTCCAAAGTC 265
Qy      137 CACCCCA 143
Db      265 CACCCCA 259

RESULT 11
AV645393/c
LOCUS      AV645393 640 bp mRNA linear EST 15-JAN-2002
DEFINITION AV645393 GLA Homo sapiens cDNA clone GLAAG10 3', mRNA sequence.
ACCESSION  AV645393
VERSION     AV645393.1 GI:9866407
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 640)
AUTHORS   Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
TITLE     Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
MEDLINE   21625106
PUBMED    11752456
COMMENT   Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES             source
    source
1..640
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="GLAAG10"
    /tissue_type="corresponding non cancerous liver tissue"
    /dev_stage="Adult"
    /lab_host="SOLR"
    /clone_lib="GLA"
    /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT  152 a 162 c 164 g 153 t 9 others
ORIGIN
Query Match      25.4%; Score 127; DB 9; Length 640;
Best Local Similarity 100.0%; Pred. No. 6.8e-58;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCAGTGTGATCGG 76
Db      407 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCAGTGTGATCGG 348
Qy      77 TTTTGGCAGTACATCAATGGCGGTGATAGCGTTTGACTCAGCGGGATTTCCAAAGTC 136
Db      347 TTTTGGCAGTACATCAATGGCGGTGATAGCGTTTGACTCAGCGGGATTTCCAAAGTC 288
Qy      137 CACCCCA 143
Db      287 CACCCCA 281

RESULT 12
AV681492/c
LOCUS      AV681492 648 bp mRNA linear EST 16-JAN-2002
DEFINITION AV681492 GLA Homo sapiens cDNA clone GKAAAD01 5', mRNA sequence.
ACCESSION  AV681492
VERSION     AV681492.1 GI:10283355
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 648)
AUTHORS   Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
TITLE     Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
MEDLINE   21625106
PUBMED    11752456
COMMENT   Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES             source
    source
1..648
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="GKAAAD01"
    /tissue_type="hepatocellular carcinoma"
    /dev_stage="Adult"
    /lab_host="SOLR"
    /clone_lib="GKA"
    /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT  152 a 166 c 166 g 158 t 6 others
ORIGIN
Query Match      25.4%; Score 127; DB 9; Length 648;
Best Local Similarity 100.0%; Pred. No. 6.8e-58;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCAGTGTGATCGG 76
Db      387 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCAGTGTGATCGG 328
Qy      77 TTTTGGCAGTACATCAATGGCGGTGATAGCGTTTGACTCAGCGGGATTTCCAAAGTC 136
Db      327 TTTTGGCAGTACATCAATGGCGGTGATAGCGTTTGACTCAGCGGGATTTCCAAAGTC 268
Qy      137 CACCCCA 143
Db      267 CACCCCA 261

RESULT 13
AV681500
LOCUS      AV681500 655 bp mRNA linear EST 16-JAN-2002
DEFINITION AV681500 GLA Homo sapiens cDNA clone GKAAAD08 5', mRNA sequence.
ACCESSION  AV681500
VERSION     AV681500.1 GI:10283363
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE
1 (bases 1 to 655)
AUTHORS
Xiao,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
TITLE
Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
MEDLINE
21625106
PUBMED
11752456
COMMENT
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
Location/Qualifiers
source
1..655
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="GKAPD08"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="GKA"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT
169 a 174 c 163 g 148 t 1 others
ORIGIN
Query Match 25.4%; Score 127; DB 9; Length 655;
Best Local Similarity 100.0%; Pred. No. 6.8e-58;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTATGATCGTATTACCATGGTGATCGG 76
Db 256 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTATGATCGTATTACCATGGTGATCGG 315
Qy 77 TTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCAGCGGGATTCCAAAGTCTC 136
Db 316 TTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCAGCGGGATTCCAAAGTCTC 375
Qy 137 CACCCCA 143
Db 376 CACCCCA 382
RESULT 14
A1815377/c
LOCUS
657 bp mRNA linear EST 09-JUL-1999
DEFINITION
B0283860 Zebrafish Kidney cDNA random primed, RZPD library no: 576
Danio rerio cDNA clone RZPD clone CHBOP576F21214Q3, mRNA sequence.
ACCESSION
A1815377
VERSION
A1815377.1 GI:5430923
KEYWORDS
EST.
SOURCE
Danio rerio (zebrafish)
ORGANISM
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Osteiophysi; Cypriniformes
; Cyprinidae; Danio.
1 (bases 1 to 657)
Look.A.T. and Holloway,M.
Zebrafish Kidney cDNA
Unpublished
Contact: Thomas Look
Dana-Farber Cancer Institute
Pediatric Oncology Department, Mayer Building-630, 44 Binney Street
, Boston, MA 02115, USA
Tel: (617) 632-5826

```

```

Fax: (617) 632-6989
Email: thomas.look@dfci.harvard.edu
The clone was obtained from an RZPD array made from an adult kidney
cDNA random primed library from Leonard Zon. RZPD library number
576. This clone is available from the Resource Centre/Primary
Database of the German Human Genome Project: RZPD (Ressourcen
Zentrum Primär Datenbank). http://www.rzpd.de
Seq primer: T7.
FEATURES
Location/Qualifiers
source
1..657
/organism="Danio rerio"
/mol_type="rRNA"
/db_xref="taxon:7955"
/clone="RZPD clone CHBOP576F21214Q3"
/dev_stage="adult"
/clone_lib="Zebrafish Kidney cDNA random primed, RZPD
library no: 576"
/note="Organ: kidney; Vector: pBK-CMV"
BASE COUNT
163 a 165 c 162 g 165 t 2 others
ORIGIN
Query Match 25.4%; Score 127; DB 9; Length 657;
Best Local Similarity 100.0%; Pred. No. 6.8e-58;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTATGATCGTATTACCATGGTGATCGG 76
Db 589 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTATGATCGTATTACCATGGTGATCGG 530
Qy 77 TTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCAGCGGGATTCCAAAGTCTC 136
Db 529 TTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGACTCAGCGGGATTCCAAAGTCTC 470
Qy 137 CACCCCA 143
Db 469 CACCCCA 463
RESULT 15
AV681509/c
LOCUS
661 bp mRNA linear EST 16-JAN-2002
DEFINITION
AV681509 GKA Homo sapiens cDNA clone GKAAAC02 5', mRNA sequence.
ACCESSION
AV681509
VERSION
AV681509.1 GI:10283372
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 661)
Xiao,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
MEDLINE
21625106
PUBMED
11752456
COMMENT
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
Location/Qualifiers
source
1..661
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"

```



```
/clone="GKAAAC02"  
/tissue type="hepatocellular carcinoma"  
/dev stage="Adult"  
/lab_host="SOLR"  
/clone_lib="GKA"  
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:  
XhoI"
```

BASE COUNT 160 a 170 c 166 g 165 t

ORIGIN

```
Query Match 25.4%; Score 127; DB 9; Length 661;  
Best Local Similarity 100.0%; Pred. No. 6.9e-58;  
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 17 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGCATCGG 76  
DB 422 GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGCATCGG 363  
  
QY 77 TTTTGGCAGTACATCAATGGGGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTC 136  
DB 362 TTTTGGCAGTACATCAATGGGGGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTC 303  
  
QY 137 CACCCCA 143  
DB 302 CACCCCA 296
```

Search completed: February 16, 2004, 03:15:41
Job time : 1319.95 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 19:10:26 ; Search time 3961.99 Seconds
(without alignments)
10335.846 Million cell updates/sec

Title: US-09-921-143-36_COPY_2000_3000

Perfect score: 1001

Sequence: 1 tctcaggtaccagtgagg.....ttcggtgcggcagcggtta 1001

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 2045481386 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.ox.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg_hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htg_mus.*
- 34: em.htg_pln.*
- 35: em.htg_rod.*
- 36: em.htg_nam.*
- 37: em.htg_vit.*
- 38: em.sy.*
- 39: em.htgo_hum.*
- 40: em.htgo_mus.*
- 41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	679	67.8	278936	2	AC098563	AC098563 Rattus no
2	679	67.8	288862	2	AC099449	AC099449 Rattus no
3	494	49.4	1146	10	RNINS2	V01243 Rat Gene fo
4	358	35.8	3979	6	A30511	A30511 Plasmid p26
5	358	35.8	3980	6	A30509	A30509 Plasmid p26
6	343	34.3	2852	10	RATINSII	J00748 Rat insulin
7	291	29.1	9632	6	BD085143	BD085143 Ricin-lik
8	290	29.0	450	6	E33315	E33315 Compound al
9	290	29.0	584	6	AR272910	AR272910 Sequence
10	290	29.0	584	6	AR276491	AR276491 Sequence
11	290	29.0	584	6	AX063012	AX063012 Sequence
12	290	29.0	584	6	AX367929	AX367929 Sequence
13	290	29.0	655	6	AX351379	AX351379 Sequence
14	290	29.0	701	6	AR072501	AR072501 Sequence
15	290	29.0	701	6	AR274302	AR274302 Sequence
16	290	29.0	713	6	AR072503	AR072503 Sequence
17	290	29.0	713	6	AR274304	AR274304 Sequence
18	290	29.0	716	6	A85396	A85396 Sequence 55
19	290	29.0	716	6	AR154889	AR154889 Sequence
20	290	29.0	716	6	E65414	E65414 Genome DNA
21	290	29.0	717	6	AR072502	AR072502 Sequence
22	290	29.0	717	6	AR274303	AR274303 Sequence
23	290	29.0	734	4	AF261959	AF261959 Anourosor
24	290	29.0	738	6	AR036905	AR036905 Sequence
25	290	29.0	738	6	AR141144	AR141144 Sequence
26	290	29.0	738	6	AR181919	AR181919 Sequence
27	290	29.0	739	4	AF261960	AF261960 Anourosor
28	290	29.0	742	6	AX197652	AX197652 Sequence
29	290	29.0	759	8	BN4489202	BN4489202 Beta nana
30	290	29.0	768	4	AF261962	AF261962 Anourosor
31	290	29.0	770	6	BD124786	BD124786 Primer fo
32	290	29.0	770	6	BD126222	BD126222 Primer fo
33	290	29.0	788	4	AF261963	AF261963 Anourosor
34	290	29.0	803	6	AX283744	AX283744 Sequence
35	290	29.0	809	6	AR244231	AR244231 Sequence
36	290	29.0	834	8	BN4489201	BN4489201 Beta nana
37	290	29.0	848	6	AR244167	AR244167 Sequence
38	290	29.0	910	9	LDO244003	LDO244003 Lepilemur
39	290	29.0	972	6	AX046223	AX046223 Sequence
40	290	29.0	1024	6	AR244042	AR244042 Sequence
41	290	29.0	1738	6	AR022240	AR022240 Sequence
42	290	29.0	1738	6	AR031416	AR031416 Sequence
43	290	29.0	1738	6	AR072989	AR072989 Sequence
44	290	29.0	1738	6	AR078254	AR078254 Sequence
45	290	29.0	1738	6	AR154630	AR154630 Sequence

ALIGNMENTS

RESULT 1
AC098563 AC098563 278936 bp DNA linear HTG 13-MAY-2003
LOCUS Rattus norvegicus clone CH230-123A15, WORKING DRAFT SEQUENCE.
DEFINITION AC098563
ACCESSION AC098563
VERSION AC098563.5 GI:30581471
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 278936)
Muzny,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsebrooks,S., Amin,A., Anguiano,D.,

* as soon as it is available and the accession number will
* be preserved.

1 281836: contig of 281836 bp in length
281837 281936: gap of unknown length
281937 282992: contig of 1056 bp in length
282993 283092: gap of unknown length
283093 284135: contig of 1043 bp in length
284136 284236: gap of unknown length
284237 285379: contig of 1144 bp in length
285380 285479: gap of unknown length
285480 286963: contig of 1484 bp in length
286964 287063: gap of unknown length
287064 288862: contig of 1799 bp in length.

FEATURES

source
1. 288862
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-57B2"
misc_feature
1. 2458
/note="wgs contig"
misc_feature
complement(280810..281268)
/note="clone boundary
clone end:Sp5
site:ECORI
end sequence: BH296258"
BASE COUNT 75546 a 55666 c 58691 g 65812 t 33147 others
ORIGIN

Query Match 67.8%; Score 679; DB 2; Length 288862;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 679; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TCTCAGGCTAACCAAGTGGGGGACATTTCTGTGGCAGCTAGACATATGTAACATGGT 60
Db 226123 TCTCAGGCTAACCAAGTGGGGGACATTTCTGTGGCAGCTAGACATATGTAACATGGT 226182
Qy 61 AGCTGCCAGAGAGAGTGAATCTCTTAACTAGTCTCTAGTGGTGAACGGTGGCTAG 120
Db 226183 AGCTGCCAGAGAGAGTGAATCTCTTAACTAGTCTCTAGTGGTGAACGGTGGCTAG 226242
Qy 121 GCCCCAGGATAGTACTATTGTTGGGACCCCATAGAGCACTGCACTGATGAGGGATGT 180
Db 226243 GCCCCAGGATAGTACTATTGTTGGGACCCCATAGAGCACTGCACTGATGAGGGATGT 226302
Qy 181 AACAGGATGTAGTTTGGAGGCCCATATGCCATTATGACCACTGACTTGTCTCAC 240
Db 226303 AACAGGATGTAGTTTGGAGGCCCATATGCCATTATGACCACTGACTTGTCTCAC 226362
Qy 241 AGCCATGCAACCCCTGCTCTCTGCTGCTGACTTAGCAGGGGATAAAGTGAGAGAAAGCCTG 300
Db 226363 AGCCATGCAACCCCTGCTCTCTGCTGCTGACTTAGCAGGGGATAAAGTGAGAGAAAGCCTG 226422
Qy 301 GGCTAATCAGGGGTGCTCAGCTCTCTCTAACTGATTTGCTCTATGTTCTTTGTTCT 360
Db 226423 GGCTAATCAGGGGTGCTCAGCTCTCTCTAACTGATTTGCTCTATGTTCTTTGTTCT 226482
Qy 361 GTGCTGCTGATGCTCTGCCCTGTGCTGACATGACCTCCCTGGCAGTGGCACAACCTGGAGC 420
Db 226483 GTGCTGCTGATGCTCTGCCCTGTGCTGACATGACCTCCCTGGCAGTGGCACAACCTGGAGC 226542
Qy 421 TGGGTGAGGCGCGGGGCGAGTGAACCTTCAGACCTTGGCACTGGAGGTGGCCCGGCAGA 480
Db 226543 TGGGTGAGGCGCGGGGCGAGTGAACCTTCAGACCTTGGCACTGGAGGTGGCCCGGCAGA 226602
Qy 481 AGCCGCGCATGCTGGATCAGTGTGTCACCAAGCATCTCTCTCTCTACCACTGGAGAACT 540
Db 226603 AGCCGCGCATGCTGGATCAGTGTGTCACCAAGCATCTCTCTCTCTACCACTGGAGAACT 226662
Qy 541 ACTGCAACTAGGCGCCACCACTACCTCTGCCCCCTCTGCAATGAATAAACCCTTTGAAA 600
Db 226663 ACTGCAACTAGGCGCCACCACTACCTCTGCCCCCTCTGCAATGAATAAACCCTTTGAAA 226722

Qy 601 GAGCACTACAAAGTTGTGTACATGCTGCATGTGCATATGTGTGTGCGGGGAAACATGA 660
Db 226723 GAGCACTACAAAGTTGTGTACATGCTGCATGTGCATATGTGTGTGCGGGGAAACATGA 226782
Qy 661 GTGGGGCTGGCTGGAGTGG 679
Db 226783 GTGGGGCTGGCTGGAGTGG 226801
RESULT 3
RNINS2
LOCUS
DEFINITION Rat gene for insulin 2.
ACCESSION V01243
VERSION V01243.1 GI:56489
KEYWORDS insulin; preproinsulin.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1146)
AUTHORS Lomedico P., Rosenthal N., Efstratidis A., Gilbert W.,
Kolodner R. and Tizard R.
TITLE The structure and evolution of the two nonallelic rat preproinsulin
genes
JOURNAL Cell 18 (2), 545-558 (1979)
MEDLINE 80045035
PUBMED 498284
REFERENCE 2
AUTHORS Kozak M.
TITLE Point mutations close to the AUG initiator codon affect the
efficiency of translation of rat preproinsulin in vivo
JOURNAL Nature 308 (5956), 241-246 (1984)
MEDLINE 84142290
PUBMED 6700727
FEATURES
Location/Qualifiers
1. 1146
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
prim_transcript 43..1103
exon 43..82
/number=1
intron 83..201
/number=1
exon 202..404
/number=2
CDS join(219..404,904..1050)
/codon_start=1
/product="preproinsulin"
/protein_id="CA24560.1"
/db_xref="GI:56490"
/db_xref="SWISS-PROT:P01323"
/translation="MALWIFLELLALLIWEPRPAQAFVKHLCQSHLVEALYLVG
ERGFYTPMGRREVDEPQVAQLGLGGPGAGDLQTLALEVARQKRGIVDQCCTSI
YQLENYCN"
intron 405..903
/number=2
exon 904..1103
/number=3
BASE COUNT 236 a 313 c 308 g 289 t
ORIGIN

Query Match 49.4%; Score 494; DB 10; Length 1146;
Best Local Similarity 99.58; Pred. No. 6.4e-278;
Matches 644; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TCTCAGGCTAACCAAGTGGGGGACATTTCTGTGGCAGCTAGACATATGTAACATGGT 60
Db 500 TCTCAGGCTAACCAAGTGGGGGACATTTCTGTGGCAGCTAGACATATGTAACATGGT 559

```
QY 61 AGCTGCCAGGAAGAGTGAATCTCTCTTAAGTCTCTAGTGGTGACGGGTGGCTAG 120
DB 560 AGCTGCCAAGAGGAGTGAGATCTTCTTAACTCTCTAGTGGTGACGGGTGGCTAG 619
QY 121 GCCCAGGATAGGTACCTATTATTTGGGAGCCCATAGAGCACTGCATGACTGAGGGATGGT 180
DB 620 GCCCAGGATAGGTACCTATTATTTGGGAGCCCATAGAGCACTGCATGACTGAGGGATGGT 679
QY 181 AACAGGATGTAGTATTTTGGAGGGCCCATATGTCATTGACAGGATGCTTGTCTCAC 240
DB 680 AACAGGATGTAGTATTTTGGAGGGCCCATATGTCATTGACAGGATGCTTGTCTCAC 739
QY 241 AGCCATGCAACCTTGGCTCTCTGCTGACCTTAGCAGGGATAAAGTACAGAGAAGCCCTG 300
DB 740 AGCCATGCAACCTTGGCTCTCTGCTGACCTTAGCAGGGATAAAGTACAGAGAAGCCCTG 799
QY 301 GGTCTAATCAGGGGGTGGCTCAGCTCTCTTAAGTGGTCTCTATGCTTCTTCTTCT 360
DB 800 GGTCTAATCAGGGGGTGGCTCAGCTCTCTTAAGTGGTCTCTATGCTTCTTCTTCT 859
QY 361 GTGCTGCTGATGCTTGGCTCTCTGCTGATGATGATGATGATGATGATGATGATGATG 420
DB 860 GTGCTGCTGATGCTTGGCTCTCTGCTGATGATGATGATGATGATGATGATGATGATG 919
QY 421 TGCGTGGAGCCCGGGGGCGGTGACCTTTCAGACCTTGGCACTGGAGGTGGCGGCGAGA 480
DB 920 TGCGTGGAGCCCGGGGGCGGTGACCTTTCAGACCTTGGCACTGGAGGTGGCGGCGAGA 979
QY 481 AGCGGGGATCTGGATCAGTCTGCTGACCCAGCATCTGCTCTCTTACCAACTGGAGAACT 540
DB 980 AGCGGGGATCTGGATCAGTCTGCTGACCCAGCATCTGCTCTCTTACCAACTGGAGAACT 1039
QY 541 ACTGCACTAGGCCACCACTACCTCTGCTGATGATGATGATGATGATGATGATGATGATA 600
DB 1040 ACTGCACTAGGCCACCACTACCTCTGCTGATGATGATGATGATGATGATGATGATA 1099
QY 601 GAGCACTACAAGTGTGTATGATGCTGCTGATGCTGATGCTGATGCTGATGCTGATG 647
DB 1100 GAGCACTACAAGTGTGTATGATGCTGCTGATGCTGATGCTGATGCTGATGCTGATG 1146

RESULT 4
A30511
LOCUS A30511 3979 bp DNA linear PAT 05-JUL-2002
DEFINITION Plasmid p264BAM.
ACCESSION A30511
VERSION A30511.1 GI:23957137
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM Fountoulakis, M., Garotta, G. and Stueber, D.
REFERENCE 1 (bases 1 to 3979)
AUTHORS Fountoulakis, M., Garotta, G. and Stueber, D.
TITLE Soluble interferon-gamma receptors and methods for their production
JOURNAL Patent: EP 039502-A 78 24-OCT-1990;
F. HOFFMANN-LA ROCHE AG
FEATURES
    source
        1..3979
            /organism="synthetic construct"
            /mol_type="genomic DNA"
            /db_xref="taxon:32630"
BASE COUNT 946 a 988 c 1115 g 930 t
ORIGIN
    Query Match 35.8%; Score 358; DB 6; Length 3979;
    Best Local Similarity 99.6%; Pred. No. 5e-198;
    Matches 458; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 171 GAGGGATGTTAACAGGATGTAGTATTTGGAGGCCCATATGTCATTGACCAAGTGA 230
DB 873 GAGGGATGTTAACAGGATGTAGTATTTGGAGGCCCATATGTCATTGACCAAGTGA 932
QY 231 CTTGTCTCAGAGCATGCAACCCCTTGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 290
```

```
DB 933 CTTGTCTCAGAGCATGCAACCCCTTGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 992
QY 291 AGAAAGCTGGGCTTAATCAGGGGGTGCCTCAGCTCCTCTTAATGATTTGCTATGTGT 350
DB 993 AGAAAGCTGGGCTTAATCAGGGGGTGCCTCAGCTCCTCTTAATGATTTGCTATGTGT 1052
QY 351 CTTTGTCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 410
DB 1053 CTTTGTCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1112
QY 411 CAATCGAGCTGGGTGGAGCCCGGGGCGAGTGCCTTCAGACCTTCGACCTGCGAGGTG 470
DB 1113 CAATCGAGCTGGGTGGAGCCCGGGGCGAGTGCCTTCAGACCTTCGAGCTTGGCAGTGGAGGTG 1172
QY 471 GCCCGGCGAAGCGCGGCATCGTGGATCAGTGTGTCACCCAGCATCTCTCTCTACCAA 530
DB 1173 GCCCGGCGAAGCGCGGCATCGTGGATCAGTGTGTCACCCAGCATCTCTCTCTACCAA 1232
QY 531 CTGGAGAACTACTGCAACTAGGCCACCACTACCTGTCACCCCTTCGCAATGATATAA 590
DB 1233 CTGGAGAACTACTGCAACTAGGCCACCACTACCTGTCACCCCTTCGCAATGATATAA 1292
QY 591 ACCTTTGAAAGAGCACTACAAAGTGTGTATCATGCGTGC 630
DB 1293 ACCTTTGAAAGAGCACTACAAAGTGTGTATCATGCGTGC 1332

RESULT 5
A30509
LOCUS A30509 3980 bp DNA linear PAT 05-JUL-2002
DEFINITION Plasmid p264BGL.
ACCESSION A30509
VERSION A30509.1 GI:23957136
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM Fountoulakis, M., Garotta, G. and Stueber, D.
REFERENCE 1 (bases 1 to 3980)
AUTHORS Fountoulakis, M., Garotta, G. and Stueber, D.
TITLE Soluble interferon-gamma receptors and methods for their production
JOURNAL Patent: EP 0393502-A 76 24-OCT-1990;
F. HOFFMANN-LA ROCHE AG
FEATURES
    source
        1..3980
            /organism="synthetic construct"
            /mol_type="genomic DNA"
            /db_xref="taxon:32630"
BASE COUNT 947 a 987 c 1114 g 931 t 1 others
ORIGIN
    Query Match 35.8%; Score 358; DB 6; Length 3980;
    Best Local Similarity 99.6%; Pred. No. 5e-198;
    Matches 458; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 171 GAGGGATGTTAACAGGATGTAGTATTTGGAGGCCCATATGTCATTGACCAAGTGA 230
DB 874 GAGGGATGTTAACAGGATGTAGTATTTGGAGGCCCATATGTCATTGACCAAGTGA 933
QY 231 CTTGTCTCAGAGCATGCAACCCCTTGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 290
DB 934 CTTGTCTCAGAGCATGCAACCCCTTGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 993
QY 291 AGAAAGCTGGGCTTAATCAGGGGGTGCCTCAGCTCCTCTCTTAATGATTTGCTATGTGT 350
DB 994 AGAAAGCTGGGCTTAATCAGGGGGTGCCTCAGCTCCTCTCTTAATGATTTGCTATGTGT 1053
QY 351 CTTTGTCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 410
DB 1054 CTTTGTCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1113
QY 411 CAATCGAGCTGGGTGGAGCCCGGGGCGAGTGCCTTCAGACCTTCGACCTTGGCAGTGGAGGTG 470
```


ORGANISM	unclassified	unclassified
REFERENCE	1 (bases 1 to 450)	
AUTHORS	Sugiyama,H., To,S. and Saito,A.	
TITLE	Compound alkylating specific base sequence in DNA and process for synthesizing the same	
JOURNAL	Patent: JP 2000159768-A 2 13-JUN-2000;	
COMMENT	SCIENCE & TECH AGENCY	
OS	PUC 18	
PN	JP 2000159768-A/2	
PD	13-JUN-2000	
PF	14-SEP-1998 JP 1998260710	
PR		
PI	HIROSHI SUGIYAMA,SHIFUKU TO,AKIRA SAITO	
PC	C07D487/04,A61K48/00,C12N15/09,C12N15/00	
CC		
Key	Location/Qualifiers	
FX	1..450	
FT	1..450	
FEATURES	Location/Qualifiers	
source	1..450	
	/organism="unidentified"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:32644"	
BASE COUNT	107 a 119 c 124 g 100 t	
ORIGIN		
Query Match	29.0%; Score 290; DB 6; Length 450;	
Best Local Similarity	100.0%; Pred.No. 4.6e-158;	
Matches 290; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	712 CGTAATCATGGTCATAGCTGTTTCTCTGTGTGAAATTGTTATCCGCTCACAAATCCACACA 771	
Db	78 CGTAATCATGGTCATAGCTGTTTCTCTGTGTGAAATTGTTATCCGCTCACAAATCCACACA 137	
QY	772 ACATACGAGCGGAACATATAAGTGTAAAGCTTGGGTGCCTATAGTAGTGAAGTAACTCA 831	
Db	138 ACATACGAGCGGAACATATAAGTGTAAAGCTTGGGTGCCTATAGTAGTGAAGTAACTCA 197	
QY	832 CATTAAATGCGTTGCGCTCACTGCCCGCTTTCACAGTCGGGAAACCTGTCTGCGCAGCTGC 891	
Db	198 CATTAAATGCGTTGCGCTCACTGCCCGCTTTCACAGTCGGGAAACCTGTCTGCGCAGCTGC 257	
QY	892 ATTAATGAATGCGCCAAACGCGGCGGAGAGCGGTTTCCGCTATTGGGGCGCTCTCCGCTT 951	
Db	258 ATTAATGAATGCGCCAAACGCGGCGGAGAGCGGTTTCCGCTATTGGGGCGCTCTCCGCTT 317	
QY	952 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTA 1001	
Db	318 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTA 367	
RESULT 9		
LOCUS	AR272910	
DEFINITION	Sequence 639 from patent US 6504010.	
ACCESSION	AR272910	
VERSION	AR272910.1 GI:29704795	
KEYWORDS		
SOURCE	Unknown.	
ORGANISM	Unknown.	
REFERENCE	1 (bases 1 to 584)	
AUTHORS	Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S.,	
TITLE	Carter,D., Retter,M.W., Mannion,J. and Fan,L.	
JOURNAL	Compositions and methods for the therapy and diagnosis of lung cancer	
FEATURES	Patent: US 6504010-A 639 07-JAN-2003;	
source	Location/Qualifiers	
	1..584	
	/organism="unknown"	
BASE COUNT	142 a 159 c 153 g 130 t	
ORIGIN		
Query Match	29.0%; Score 290; DB 6; Length 450;	
Best Local Similarity	100.0%; Pred.No. 4.6e-158;	
Matches 290; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	712 CGTAATCATGGTCATAGCTGTTTCTCTGTGTGAAATTGTTATCCGCTCACAAATCCACACA 771	
Db	78 CGTAATCATGGTCATAGCTGTTTCTCTGTGTGAAATTGTTATCCGCTCACAAATCCACACA 137	
QY	772 ACATACGAGCGGAACATATAAGTGTAAAGCTTGGGTGCCTATAGTAGTGAAGTAACTCA 831	
Db	138 ACATACGAGCGGAACATATAAGTGTAAAGCTTGGGTGCCTATAGTAGTGAAGTAACTCA 197	
QY	832 CATTAAATGCGTTGCGCTCACTGCCCGCTTTCACAGTCGGGAAACCTGTCTGCGCAGCTGC 891	
Db	198 CATTAAATGCGTTGCGCTCACTGCCCGCTTTCACAGTCGGGAAACCTGTCTGCGCAGCTGC 257	
QY	892 ATTAATGAATGCGCCAAACGCGGCGGAGAGCGGTTTCCGCTATTGGGGCGCTCTCCGCTT 951	
Db	258 ATTAATGAATGCGCCAAACGCGGCGGAGAGCGGTTTCCGCTATTGGGGCGCTCTCCGCTT 317	
QY	952 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTA 1001	
Db	318 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTA 367	
RESULT 9		
LOCUS	AR272910	
DEFINITION	Sequence 639 from patent US 6504010.	
ACCESSION	AR272910	
VERSION	AR272910.1 GI:29704795	
KEYWORDS		
SOURCE	Unknown.	
ORGANISM	Unknown.	
REFERENCE	1 (bases 1 to 584)	
AUTHORS	Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S.,	
TITLE	Carter,D., Retter,M.W., Mannion,J. and Fan,L.	
JOURNAL	Compositions and methods for the therapy and diagnosis of lung cancer	
FEATURES	Patent: US 6504010-A 639 07-JAN-2003;	
source	Location/Qualifiers	
	1..584	
	/organism="unknown"	
BASE COUNT	142 a 159 c 153 g 130 t	
ORIGIN		

Query Match 29.0%; Score 290; DB 6; Length 584;
 Best Local Similarity 100.0%; Pred. No. 4.6e-158; Indels 0; Gaps 0;
 Matches 290; Conservative 0; Mismatches 0;

QY 712 CGTAATCATGTCATAGCTGTTTCTCTGTGTGAATTTATCCGCTCAAAATCCACACA 771
 Db 37 CGTAATCATGTCATAGCTGTTTCTCTGTGTGAATTTATCCGCTCAAAATCCACACA 96

QY 772 ACATACGAGCGGAGCATAAAGTCTAAAGCTGGGTGGCTTAATGAGTGAGCTAACTCA 831
 Db 97 ACATACGAGCGGAGCATAAAGTCTAAAGCTGGGTGGCTTAATGAGTGAGCTAACTCA 156

QY 832 CATTAAATTCGGTTCGCTCACTGCCGCTTTCAGTCGCGGAACCTGTCGCCAGTGC 891
 Db 157 CATTAAATTCGGTTCGCTCACTGCCGCTTTCAGTCGCGGAACCTGTCGCCAGTGC 216

QY 892 ATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATTGGCGCTCTCCGCTT 951
 Db 217 ATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATTGGCGCTCTCCGCTT 276

QY 952 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTA 1001
 Db 277 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTA 326

RESULT 10
 LOCUS AR276491 584 bp DNA linear PAT 10-APR-2003
 DEFINITION Sequence 639 from patent US 6509448.
 ACCESSION AR276491
 VERSION AR276491.1 GI:29710136
 KEYWORDS Unknown.
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 584)
 AUTHORS Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S., Carter, D., Retter, M.W., Mammion, J., Pan, L., and Wang, A.
 TITLE Compositions and methods for the therapy and diagnosis of lung cancer
 JOURNAL Patent: US 6509448-A 639 21-JAN-2003;
 FEATURES Location/Qualifiers
 source 1..584 /organism="unknown"
 BASE COUNT 142 a 153 c 130 t
 ORIGIN

Query Match 29.0%; Score 290; DB 6; Length 584;
 Best Local Similarity 100.0%; Pred. No. 4.6e-158; Indels 0; Gaps 0;
 Matches 290; Conservative 0; Mismatches 0;

QY 712 CGTAATCATGTCATAGCTGTTTCTCTGTGTGAATTTATCCGCTCAAAATCCACACA 771
 Db 37 CGTAATCATGTCATAGCTGTTTCTCTGTGTGAATTTATCCGCTCAAAATCCACACA 96

QY 772 ACATACGAGCGGAGCATAAAGTCTAAAGCTGGGTGGCTTAATGAGTGAGCTAACTCA 831
 Db 97 ACATACGAGCGGAGCATAAAGTCTAAAGCTGGGTGGCTTAATGAGTGAGCTAACTCA 156

QY 832 CATTAAATTCGGTTCGCTCACTGCCGCTTTCAGTCGCGGAACCTGTCGCCAGTGC 891
 Db 157 CATTAAATTCGGTTCGCTCACTGCCGCTTTCAGTCGCGGAACCTGTCGCCAGTGC 216

QY 892 ATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATTGGCGCTCTCCGCTT 951
 Db 217 ATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATTGGCGCTCTCCGCTT 276

QY 952 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTA 1001
 Db 277 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTA 326

RESULT 11
 LOCUS AX063012 584 bp DNA linear PAT 24-JAN-2001
 DEFINITION Sequence 639 from Patent WO0100828.
 ACCESSION AX063012
 VERSION AX063012.1 GI:12540870
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 REFERENCE 1
 AUTHORS Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S., Carter, D., Retter, M.W., and Mammion, J.
 TITLE Compositions and methods for the therapy and diagnosis of lung cancer
 JOURNAL Patent: WO 0100828-A 639 04-JAN-2001;
 FEATURES CORIXA CORPORATION (US)
 source Location/Qualifiers
 1..584 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 142 a 153 c 130 t
 ORIGIN

Query Match 29.0%; Score 290; DB 6; Length 584;
 Best Local Similarity 100.0%; Pred. No. 4.6e-158; Indels 0; Gaps 0;
 Matches 290; Conservative 0; Mismatches 0;

QY 712 CGTAATCATGTCATAGCTGTTTCTCTGTGTGAATTTATCCGCTCAAAATCCACACA 771
 Db 37 CGTAATCATGTCATAGCTGTTTCTCTGTGTGAATTTATCCGCTCAAAATCCACACA 96

QY 772 ACATACGAGCGGAGCATAAAGTCTAAAGCTGGGTGGCTTAATGAGTGAGCTAACTCA 831
 Db 97 ACATACGAGCGGAGCATAAAGTCTAAAGCTGGGTGGCTTAATGAGTGAGCTAACTCA 156

QY 832 CATTAAATTCGGTTCGCTCACTGCCGCTTTCAGTCGCGGAACCTGTCGCCAGTGC 891
 Db 157 CATTAAATTCGGTTCGCTCACTGCCGCTTTCAGTCGCGGAACCTGTCGCCAGTGC 216

QY 892 ATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATTGGCGCTCTCCGCTT 951
 Db 217 ATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATTGGCGCTCTCCGCTT 276

QY 952 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTA 1001
 Db 277 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTA 326

RESULT 12
 LOCUS AX367929 584 bp DNA linear PAT 16-FEB-2002
 DEFINITION Sequence 639 from Patent WO0204514.
 ACCESSION AX367929
 VERSION AX367929.1 GI:18856016
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 REFERENCE 1
 AUTHORS Wang, T., Watanabe, Y., Henderson, R.A., Johnson, J.C., Retter, M.W., Marnerakis, M., Carter, D., Fanger, G.R., Vedvick, T.S., Bangur, C.S., Mcnabb, A., Fanger, N., Switzer, A., McNeill, P.D., and Clapper, J.D.
 TITLE Compositions and methods for the therapy and diagnosis of lung cancer
 JOURNAL Patent: WO 0204514-A 639 17-JAN-2002;
 FEATURES CORIXA CORPORATION (US)
 source Location/Qualifiers
 1..584 /organism="Homo sapiens"

Query Match 29.0%; Score 290; DB 6; Length 584;
 Best Local Similarity 100.0%; Pred. No. 4.6e-158; Indels 0; Gaps 0;
 Matches 290; Conservative 0; Mismatches 0;

QY 712 CGTAATCATGTCATAGCTGTTTCTCTGTGTGAATTTATCCGCTCAAAATCCACACA 771
 Db 37 CGTAATCATGTCATAGCTGTTTCTCTGTGTGAATTTATCCGCTCAAAATCCACACA 96

QY 772 ACATACGAGCGGAGCATAAAGTCTAAAGCTGGGTGGCTTAATGAGTGAGCTAACTCA 831
 Db 97 ACATACGAGCGGAGCATAAAGTCTAAAGCTGGGTGGCTTAATGAGTGAGCTAACTCA 156

QY 832 CATTAAATTCGGTTCGCTCACTGCCGCTTTCAGTCGCGGAACCTGTCGCCAGTGC 891
 Db 157 CATTAAATTCGGTTCGCTCACTGCCGCTTTCAGTCGCGGAACCTGTCGCCAGTGC 216

QY 892 ATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATTGGCGCTCTCCGCTT 951
 Db 217 ATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATTGGCGCTCTCCGCTT 276

QY 952 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTA 1001
 Db 277 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTA 326

Db	211	ATTAATGATCGCCACACGCGGGAGAGCGGTTTTCGGTATTTCGGCGGCTCTTCGCGCTT	270
Qy	952	CCTCGCTCACTGACTCGCTCGCTCGTTCGTTTCGGCTGGCGGAGCGGTA	1001
Db	271	CCTCGCTCACTGACTCGCTCGCTCGTTCGTTTCGGCTGGCGGAGCGGTA	320
RESULT 14			
LOCUS	AR072501	701 bp	DNA
DEFINITION	Sequence 24 from patent US 5948614.		linear
ACCESSION	AR072501		
VERSION	AR072501.1	GI:9999265	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 701)		
TITLE	Chatterjee, D.K.		
JOURNAL	Cloned DNA polymerases from thermotoga maritima and mutants thereof		
FEATURES	Patent: US 5948614-A 24 07-SEP-1999;		
source	Location/Qualifiers		
	1..701		
	/organism="unknown"		
BASE COUNT	154 a 179 c 191 g	152 t	25 others
ORIGIN			
	Query Match	29.0%;	Score 290; DB 6; Length 701;
	Best Local Similarity	100.0%;	Pred. No. 4..6e-158;
	Matches	290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	712	CCTAATCATGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACA	771
Db	47	CCTAATCATGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACA	106
Qy	772	ACATACGAGCGGAGCATAAAGTAAAGCCTGGGGTGCTTAATGAGTGAGCTAACTCA	831
Db	107	ACATACGAGCGGAGCATAAAGTAAAGCCTGGGGTGCTTAATGAGTGAGCTAACTCA	166
Qy	832	CATTAAATTCGTTGCGCTCACTGCCCGCTTTCAGTCGGGAACCTGTGTGCCAGCTGC	891
Db	167	CATTAAATTCGTTGCGCTCACTGCCCGCTTTCAGTCGGGAACCTGTGTGCCAGCTGC	226
Qy	892	ATTAAATGATCGGCCAACGCGGGGAGAGCGGTTTTCGTTATTTGGCGCTCTTTCGCGTT	951
Db	227	ATTAAATGATCGGCCAACGCGGGGAGAGCGGTTTTCGTTATTTGGCGCTCTTTCGCGTT	286
Qy	952	CCTCGCTCACTGACTCGCTCGCTCGCTCGTTCGTTTCGGCTGGCGGAGCGGTA	1001
Db	287	CCTCGCTCACTGACTCGCTCGCTCGTTCGTTTCGGCTGGCGGAGCGGTA	336
RESULT 15			
LOCUS	AR274302	701 bp	DNA
DEFINITION	Sequence 24 from patent US 6506560.		linear
ACCESSION	AR274302		
VERSION	AR274302.1	GI:29706732	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 701)		
TITLE	Hughes, A.J., and Chatterjee, D.K.		
JOURNAL	Cloned DNA polymerases from Thermotoga and mutants thereof		
FEATURES	Patent: US 6506560-A 24 14-JAN-2003;		
source	Location/Qualifiers		
	1..701		
	/organism="unknown"		
BASE COUNT	154 a 179 c 191 g	152 t	25 others
ORIGIN			
	Query Match	29.0%;	Score 290; DB 6; Length 701;

Best Local Similarity 100.0%; Pred. No. 4.6e-158;
 Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	712	CGTAATCATGTCATAGCTGTTTCCGTGTGAAATTTGTAATCCGCTCACAATCCACACA	771
Db	47	CGTAATCATGTCATAGCTGTTTCCGTGTGAAATTTGTAATCCGCTCACAATCCACACA	106
QY	772	ACATACGAGCCGGAAGCATTAAGCTGAAGCTGGGTGCTTAATGAGTGAGCTAACTCA	831
Db	107	ACATACGAGCCGGAAGCATTAAGCTGAAGCTGGGTGCTTAATGAGTGAGCTAACTCA	166
QY	832	CATTAAATTCGCTTGGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGC	891
Db	167	CATTAAATTCGCTTGGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGC	226
QY	892	ATTAAATGATCGGCCCAACGCGGGGAGAGCGGTTTTCGTATTGGGGGCTCTTCGCTT	951
Db	227	ATTAAATGATCGGCCCAACGCGGGGAGAGCGGTTTTCGTATTGGGGGCTCTTCGCTT	286
QY	952	CCTCGCTCACTGACTGCTGCTCGCTCGGTTCGTTCGGCTGCGCGGAGCGGTA	1001
Db	287	CCTCGCTCACTGACTGCTGCTCGCTCGGTTCGTTCGGCTGCGCGGAGCGGTA	336

Search completed: February 16, 2004, 01:25:51
 Job time : 3967.99 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 19:07:11 ; Search time 324.065 Seconds
(without alignments)
8338.264 Million cell updates/sec

Title: US-09-921-143-36_COPY_2000_3000
Perfect score: 1001
Sequence: 1 tctcagctaacctagtgggg.....ttcggctgaggcagcggtta 1001

Scoring table: OLIGO NUC
Gapop 60.0, Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 0
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

- Database : N_Geneseq_19Jun03.*
- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
 - 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
 - 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
 - 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
 - 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
 - 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
 - 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
 - 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
 - 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
 - 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
 - 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
 - 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
 - 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
 - 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
 - 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
 - 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
 - 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
 - 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
 - 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
 - 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
 - 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
 - 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
 - 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
 - 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
 - 25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1001	100.0	5283	ABK10062	Expression vector
2	309	30.9	3979	AA006312	Sequence of plasmid
3	307	30.7	5552	AA58309	PDIP/PyQSP.1 plasm
4	291	29.1	9632	AA97898	Baculovirus transf
5	291	29.1	9632	AA04251	Baculovirus transf
6	290	29.0	409	ABV45293	Human prostate exp
7	290	29.0	450	AAA10990	Fragment #2 of pUC
8	290	29.0	515	AAA82004	N. meningitidis pa

9	290	29.0	561	23	ABV45308	Human prostate exp
10	290	29.0	584	22	AAF68704	Human lung tumour
11	290	29.0	584	24	ABK38615	cDNA encoding clon
12	290	29.0	584	25	ACA10944	Human lung neuroen
13	290	29.0	584	25	ABX98995	Lung cancer therap
14	290	29.0	598	21	AA62566	Human breast cance
15	290	29.0	598	21	ABT12681	Orestes sequencing
16	290	29.0	655	24	ABK27689	Human colon cancer
17	290	29.0	686	22	ABA20358	Human nervous syst
18	290	29.0	686	22	ABA21448	Human nervous syst
19	290	29.0	686	22	AA105999	Human reproductive
20	290	29.0	686	22	AA107025	Human reproductive
21	290	29.0	686	22	AA527722	DNA encoding novel
22	290	29.0	686	22	AAK81476	Human immune/haema
23	290	29.0	686	22	AA533387	DNA encoding human
24	290	29.0	686	23	ABL98563	Human testicular a
25	290	29.0	686	24	ABL69910	Human secreted pro
26	290	29.0	686	25	ABZ74009	Secreted protein g
27	290	29.0	686	25	ABZ74600	Secreted protein g
28	290	29.0	686	25	ABZ68007	Human secreted pro
29	290	29.0	686	25	ABZ68121	Human secreted pro
30	290	29.0	691	21	AAA81954	N. meningitidis pa
31	290	29.0	717	21	AAZ88769	T. neapolitana Tne
32	290	29.0	721	21	AAA81967	N. meningitidis pa
33	290	29.0	728	21	AAA81752	N. meningitidis pa
34	290	29.0	738	20	AAZ03716	Plasmid pGEM DNA i
35	290	29.0	739	22	AAK61981	Human immune/haema
36	290	29.0	742	22	AAH83483	Human ovarian tumo
37	290	29.0	752	19	AAV31294	E. coli J96 pathog
38	290	29.0	770	22	AAK91757	Human CDNA 5'-end
39	290	29.0	770	22	AAK93193	Human CDNA clone r
40	290	29.0	803	24	AA561891	Porcine muscular s
41	290	29.0	809	21	AAZ97398	Human prostate can
42	290	29.0	848	21	AAZ97334	Human prostate can
43	290	29.0	972	21	AA83524	T7 promoter sequen
44	290	29.0	1024	21	AAZ97209	Human prostate can
45	290	29.0	1091	23	ABK43594	DNA encoding novel

ALIGNMENTS

RESULT 1
ABK10062
ID ABK10062 standard; DNA; 5283 BP.

XX ABK10062;

XX 21-MAY-2002 (first entry)

XX Expression vector construct pVGI.1 containing VEGF-2 insert.

XX Human; vascular endothelial growth factor 2; VEGF-2; ds; pVGI.1;
XX chronic limb ischaemia; myocardial ischaemia; autoimmune disorder; PHE4a;
XX allergic reaction; organ rejection; inflammatory condition; arrhythmia;
XX hyperproliferative disorder; viral infection; bacterial infection;
XX fungal infection; parasitic infection; cardiovascular disorder; embolism;
XX heart valve disease; aneurysm; arterial occlusive disorder; gene therapy.

XX Synthetic.

XX WO200211769-A1.

XX 14-FEB-2002.

XX 03-AUG-2001; 2001WO-US24658.

XX 04-AUG-2000; 2000US-223276P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Coleman TA;

XX

DR WPI; 2002-217153/27.

XX Isolated nucleic acid having expression vector construct with vascular
PT endothelial growth factor-2 insert, useful for treating chronic limb
PT ischaemia or myocardial ischaemia, autoimmune disorders and allergic
PT conditions -

XX Disclosure; Fig 31; 241pp; English.

XX The invention relates to an isolated nucleic acid comprising pVGI.1
CC expression vector construct containing the vascular endothelial growth
CC factor 2 (VEGF-2) insert. The nucleic acid is useful for producing a host
CC cell by transducing, transforming or transfecting a host cell with the
CC DNA and for treating a patient having chronic limb ischaemia or
CC myocardial ischaemia, or a disease or disorder selected from autoimmune
CC disorders (e.g. rheumatoid arthritis, dermatitis), allergic reactions or
CC conditions (e.g. asthma), organ rejection, inflammatory conditions (e.g.
CC Crohn's disease), hyperproliferative disorders (e.g. Gaucher's disease),
CC diseases due to viral, bacterial, fungal or parasitic infection,
CC cardiovascular disorders (e.g. cardiomyopathy), arrhythmia, heart valve
CC diseases, aneurysms, arterial occlusive disorders and embolism. This
CC sequence represents the pVGI.1 expression vector containing the VEGF-2
CC insert.

XX SQ Sequence 5283 BP; 1244 A; 1422 C; 1417 G; 1200 T; 0 Other;

Query Match 100.0%; Score 1001; DB 24; Length 5283;

Best Local Similarity 100.0%; Pred. No. 6.8e-300;

Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTCAGGCTAACCACTGGGGGACATTTCTGGGAGCTAGACATATGTAACATGGT 60
DB 2000 TCTCAGGCTAACCACTGGGGGACATTTCTGGGAGCTAGACATATGTAACATGGT 2059
QY 61 AGCTGCCAGAGAGAGTGAATCTTCTTAAAGTCTCTAGTGTGACGGGTGGCTAG 120
DB 2060 AGCTGCCAGAGAGAGTGAATCTTCTTAAAGTCTCTAGTGTGACGGGTGGCTAG 2119
QY 121 GCCCAGAGTAGGTACCTATTCTGGGAGCCCATAGAGCACTGCACCTAGAGGATGGT 180
DB 2120 GCCCAGAGTAGGTACCTATTCTGGGAGCCCATAGAGCACTGCACCTAGAGGATGGT 2179
QY 181 AACAGGATGTAGTGTGGAGCCCATATGTCATTCATGACAGTGAAGTGTCTAC 240
DB 2180 AACAGGATGTAGTGTGGAGCCCATATGTCATTCATGACAGTGAAGTGTCTAC 2239
QY 241 AGCCATGCAACCTTGCTCTGCTGCTGCTAGCACTAGCAGGGGATAAGTGAGAGAGCCTG 300
DB 2240 AGCCATGCAACCTTGCTCTGCTGCTGCTAGCACTAGCAGGGGATAAGTGAGAGAGCCTG 2299
QY 301 GCCTAATCAGGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
DB 2300 GCCTAATCAGGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2359
QY 361 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 2360 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2419
QY 421 TGGGTGAGGGCCGGGGGAGGTGACCTTCAGACCTTGGCACTGGAGGTGGCCGGCAGA 480
DB 2420 TGGGTGAGGGCCGGGGGAGGTGACCTTCAGACCTTGGCACTGGAGGTGGCCGGCAGA 2479
QY 481 AGCGCGGCACTCGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
DB 2480 AGCGCGGCACTCGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2539
QY 541 ACTGCACTAGGCCACCACTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
DB 2540 ACTGCACTAGGCCACCACTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2599
QY 601 CAGCACTACAGTGTGTGTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 2600 CAGCACTACAGTGTGTGTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2659

QY 661 GTGGGGCTGGCTGGAGTGGTGGGGCTTAATCTATCTGGGAGCTGTCTAGACGTAAATCAT 720
DB 2660 GTGGGGCTGGCTGGAGTGGTGGGGCTTAATCTATCTGGGAGCTGTCTAGACGTAAATCAT 2719
QY 721 GGTTCATAGCTGTTTCTCTGTTGAAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 2720 GGTTCATAGCTGTTTCTCTGTTGAAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2779
QY 781 CCGGAGCATTAAGTGTAAAGCTGGGTGCTTAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 840
DB 2780 CCGGAGCATTAAGTGTAAAGCTGGGTGCTTAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2839
QY 841 CGTTGGCTCTCACTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
DB 2840 CGTTGGCTCTCACTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2899
QY 901 TGGCCCAAGCCGGGGAGAGCGGTTTGGCTATTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
DB 2900 TGGCCCAAGCCGGGGAGAGCGGTTTGGCTATTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2959
QY 961 CTGACTCGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1001
DB 2960 CTGACTCGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000

RESULT 2

AAQ06312

ID AAQ06312 standard; DNA; 3979 BP.

XX AC AAQ06312;

XX DT 25-MAR-2003 (updated)

XX DT 23-JAN-1991 (first entry)

XX DE Sequence of plasmid p264BGL.

XX KW IFN-gamma receptor; autoimmune disease; multiple sclerosis;

XX KW hyper-sensitivity; ds.

XX OS Homo sapiens.

XX PH Key Location/Qualifiers

XX LTR 279..1242

XX CDS /*tag= a

XX CDS 243..426

XX CDS /*tag= b

XX CDS 243..405

XX CDS /*tag= c

XX CDS /*tag= S.P.1

XX CDS 1288..1292

XX CDS /*tag= d

XX CDS polyA_signal

XX CDS 1288..1292

XX CDS /*tag= d

XX CDS BP393502-A.

XX CDS 24-OCT-1990.

XX CDS 11-APR-1990; 90EP-0106992.

XX CDS 11-APR-1990; 90EP-0106992.

XX CDS (HOFF) HOFFMANN-LA ROCHE AG.

XX CDS Fountoulakis M, Garotta G, Stuber D;

XX CDS WPI; 1990-322042/43.

XX CDS P-PSDB; AAR07066.

XX CDS Soluble interferon-gamma receptors - for treating auto-immune

XX CDS diseases, chronic inflammations, etc.

XX CDS Disclosure; Fig 55; 174pp; English.

XX CDS

Db 7122 ACATTAATTTGGCTTGGCTCACTGCCGCTTTCCAGTCGGGAACCTGTGTGCCAGCTG 7181
QY 891 CATTAAATGAATCGGCAACGGCGGGGAGAGCGGTTTGGCTATTGGCGCTCTTCGGCT 950
Db 7182 CATTAAATGAATCGGCAACGGCGGGGAGAGCGGTTTGGCTATTGGCGCTCTTCGGCT 7241
QY 951 TCTCTCGCTCACTGACTCGCTGCGCTGCGTTCGCTGCGTTCGCTGCGGAGCGGTA 1001
Db 7242 TCTCTCGCTCACTGACTCGCTGCGCTGCGTTCGCTGCGTTCGCTGCGGAGCGGTA 7292

RESULT 6
ABV45293
ID ABV45293 standard; cDNA; 409 BP.
XX
AC ABV45293;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 45284.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX
PS Claim 1; Page 8964; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 409 BP; 78 A; 111 C; 112 G; 108 T; 0 other;

Query Match 29.0%; Score 290; DB 23; Length 409;
Best Local Similarity 100.0%; Pred. NO. 1.3e-80;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGGTCATAGCTGTTTCCGTGTGTAATTTGTTATCGGCTCACAATTCACACA 771
Db 106 CGTAATCATGGTCATAGCTGTTTCCGTGTGTAATTTGTTATCGGCTCACAATTCACACA 165
QY 772 ACATACGAGCCGGAAGCATAAAGTGTAAAGCTTGGGTGCTTAATGAGTGAGCTAACTCA 831
Db 166 ACATACGAGCCGGAAGCATAAAGTGTAAAGCTTGGGTGCTTAATGAGTGAGCTAACTCA 225
QY 832 CATTAAATGGTTGGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTGTGCCAGCTGC 891
Db 226 CATTAAATGGTTGGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTGTGCCAGCTGC 285
QY 892 ATTAATGAATCGCCAAACGGCGGGGAGAGCGGTTTGGCTATTGGCGCTCTTCGGCTT 951
Db 286 ATTAATGAATCGCCAAACGGCGGGGAGAGCGGTTTGGCTATTGGCGCTCTTCGGCTT 345
QY 952 CTTGCTCACTGACTCGCTGCGCTGCGTTCGCTGCGGAGCGGTA 1001
Db 346 CTTGCTCACTGACTCGCTGCGCTGCGTTCGCTGCGGAGCGGTA 395

RESULT 7
AAA10990
ID AAA10990 standard; DNA; 450 BP.
XX
AC AAA10990;
XX
DT 20-JUL-2000 (first entry)
XX
DE Fragment #2 of pUC 18 nucleotide sequence.
XX
KW Pyrrrole derivative; pUC 18; alkylating agent; gene expression regulator;
KW treatment; drug; cancer; ds.
XX
OS Synthetic.
XX
PN WO200015641-A1.
XX
PD 23-MAR-2000.
XX
PF 12-MAR-1999; 99WO-JP01228.
XX
PR 14-SEP-1998; 98JP-0260710.
XX
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
PI Sugiyama H, Tao ZP, Saito I;
XX WPI; 2000-271367/23.
XX

New pyrrole derivatives for alkylating specific base sequence of DNA,
useful in drug compositions for regulating expression of a gene
carrying a base sequence which can be recognized then alkylated by
them, e.g. in treatment of cancers -
PT
PT
XX
XX
PS Disclosure; Page 29; 35pp; Japanese.
XX
CC This sequence is a fragment of the pUC 18 nucleotide sequence. The
CC fragments is used in the course of the invention. The invention relates
CC to pyrrole derivatives which recognise a specific DNA base sequence, and
CC strongly bind to the adjacent base via a covalent bond. The pyrrole
CC derivatives alkylate the base sequence, once it has been recognised. The
CC invention also relates to an agent (made from the pyrrole derivatives)
CC for alkylating genes, a process for producing an agent for alkylating
CC genes, and a method for regulating the expression of a gene with a part
CC containing the specific base sequence comprising using an alkylating
CC agent. The compounds are useful in drug compositions for regulating
CC expression of a gene carrying a base sequence which can be recognized
CC then alkylated by them, e.g. in treatment of cancers.

SQ Sequence 450 BP; 107 A; 119 C; 124 G; 100 T; 0 other;
Query Match 29.0%; Score 290; DB 21; Length 450;

Best Local Similarity 100.0%; Pred. No. 1.2e-80; Mismatches 0; Indels 0; Gaps 0;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGTCATAGCTGTTTCTGTTGTAATTTGTAATTCGGCTCAAAATTCACACA 771
DB 78 CGTAATCATGTCATAGCTGTTTCTGTTGTAATTTGTAATTCGGCTCAAAATTCACACA 137

QY 772 ACATAGAGCGCGAAGCATAGTGAAGCTGGGTCCTTAATGAGTGAGCTAACTCA 831
DB 138 ACATAGAGCGCGAAGCATAGTGAAGCTGGGTCCTTAATGAGTGAGCTAACTCA 197

QY 832 CATTAATTCGTTGCGCTCACTGCCGCTTTCCAGTCGCGAAGCTGTCGCGAGCTGC 891
DB 198 CATTAATTCGTTGCGCTCACTGCCGCTTTCCAGTCGCGAAGCTGTCGCGAGCTGC 257

QY 892 ATTAATGAATCGCGCAACGCGCGGAGAGCGGTTTTCGTAATTCGGCGCTCTTCCGCTT 951
DB 258 ATTAATGAATTCGCGCAACGCGCGGAGAGCGGTTTTCGTAATTCGGCGCTCTTCCGCTT 317

QY 952 CTTGCTCATGCTCGCTGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTA 1001
DB 318 CTTGCTCATGCTCGCTGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTA 367

RESULT 8

AA82004
ID AA82004 standard; DNA; 515 BP.
XX
AC AA82004;
XX
DT 04-DEC-2000 (first entry)
XX
DE N. meningitidis partial DNA sequence gnm_551 SEQ ID NO:551.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB; ds.

OS Neisseria meningitidis.
XX
FN WO200022430-A2.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-US23573.
XX
PR 09-OCT-1998; 98US-0103794.
XX
PR 30-APR-1999; 99US-0132068.
XX

PA (CHIR) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC; Scarlato V;
PI Maignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizzi M;
XX
DR WPI; 2000-318079/27.
XX

PT Isolated nucleotide sequences of Neisseria meningitidis which can be
PT used in the diagnosis and treatment of N. meningitidis infection and
PT other Neisseria infections, for example, N.gonorrhoea
XX
PS Claim 7; Page 1653; 1760pp; English.
XX

CC The present invention describes methods of obtaining immunogenic
CC proteins from Neisseria genomic sequences. AA81453 to AA82414
CC represent specifically claimed Neisseria meningitidis genomic DNA
CC sequences; AA81260 to AA81303 and AA825620 to AA825663 represent
CC Neisseria DNA sequences and their corresponding proteins; AA81254 to
CC AA81259 and AA81304 to AA81331 represent PCR primers used in the
CC isolation of Neisseria meningitidis DNA sequences; and AA81322 to
CC AA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC sequences, which are all used in the exemplification of the present
CC invention. The nucleic acid sequences, protein sequences, and antibodies

CC against them, can be used in the manufacture of a composition. The
CC composition can be used as a medicament (or in the manufacture of a
CC medicament) for treating, preventing or diagnosing infection due to
CC Neisserial bacteria. For example, some of the identified proteins could
CC be components of vaccines against Meningococcus B; against all serotypes;
CC and/or against all pathogenic Neisseriae. Identification of sequences
CC from the bacterium will also facilitate production of biological probes,
CC particularly organism-specific probes. Attempts to make efficacious
CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC Multivalent vaccines have also been tried but none have successfully
CC overcome antigenic variability. The provision of further, complete
CC sequences may provide an opportunity to identify secreted or surface
CC exposed proteins that may be presumed targets for the immune system and
CC which are not antigenically variable or at least more conserved than
CC other more variable regions.

XX
SQ Sequence 515 BP; 110 A; 129 C; 122 G; 154 T; 0 other;

Query Match 29.0%; Score 290; DB 21; Length 515;
Best Local Similarity 100.0%; Pred. No. 1.2e-80; Indels 0; Gaps 0;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGTCATAGCTGTTTCTGTTGTAATTTGTAATTCGGCTCAAAATTCACACA 771
DB 212 CGTAATCATGTCATAGCTGTTTCTGTTGTAATTTGTAATTCGGCTCAAAATTCACACA 271

QY 772 ACATAGAGCGCGAAGCATAGTGAAGCTGGGTCCTTAATGAGTGAGCTAACTCA 831
DB 272 ACATAGAGCGCGAAGCATAGTGAAGCTGGGTCCTTAATGAGTGAGCTAACTCA 331

QY 832 CATTAATTCGTTGCGCTCACTGCCGCTTTCCAGTCGCGAAGCTGTCGCGAGCTGC 891
DB 332 CATTAATTCGTTGCGCTCACTGCCGCTTTCCAGTCGCGAAGCTGTCGCGAGCTGC 391

QY 892 ATTAATGAATCGCGCAACGCGCGGAGAGCGGTTTTCGTAATTCGGCGCTCTTCCGCTT 951
DB 392 ATTAATGAATCGCGCAACGCGCGGAGAGCGGTTTTCGTAATTCGGCGCTCTTCCGCTT 451

QY 952 CTTGCTCATGCTCGCTGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTA 1001
DB 452 CTTGCTCATGCTCGCTGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTA 501

RESULT 9

ABV45308
ID ABV45308 standard; cDNA; 561 BP.
XX
AC ABV45308;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 45299.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.

OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX

XX 17-FEB-2000; 2000US-183319P.
XX 16-MAR-2000; 2000US-189862P.
XX 25-MAY-2000; 2000US-207454P.
XX 09-JUN-2000; 2000US-211314P.
XX 18-JUL-2000; 2000US-219007P.
XX 13-DEC-2000; 2000US-255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

the polynucleotide, protein or antigen-presenting cells, under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells) and inhibiting the development of a cancer in a patient (by incubating CD4⁺ and/or CD8⁺ T cells isolated from a patient with the polynucleotide, protein or antigen presenting cells that express the polynucleotide, such that T cells proliferate, administering to the patient an effective amount of the proliferated T cells, and thus inhibiting the development of a cancer in the patient. The polynucleotide, protein and cells are useful in the treatment of a cancer, stimulating an immune response in a patient, and for treating a cancer in a patient (particularly lung cancer). The oligonucleotide is useful for determining the presence of a cancer in a patient. The protein and oligonucleotides are useful in pharmaceutical compositions, e.g. vaccines. The polynucleotide is also useful as a probe or primer for nucleic acid hybridisation, and in the design and preparation of ribozyme molecules for inhibiting expression of tumour polypeptides and proteins in tumour cells. An amplified portion of the polynucleotide is useful for isolating a full-length gene from a suitable library. The present sequence is a cDNA (full length, extended or partial) isolated from a library derived from lung tumour/cancer cells. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the USPTO at seqdata.uspto.gov/sequence.html?DocId=20020197669.

Sequence 584 BP; 142 A; 159 C; 153 G; 130 T; 0 other;

Query Match 29.0%; Score 290; DB 25; Length 584;
Best Local Similarity 100.0%; Pred. No. 1.1e-80;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGTCATAGCTGTTTCTGTCGTAATTTGTTATCCGCTCACAAATCCACACA 771
Db 37 CGTAATCATGTCATAGCTGTTTCTGTCGTAATTTGTTATCCGCTCACAAATCCACACA 96

QY 772 ACATACGAGCCGGAAGCATAAAGTGAAGCTGGGTGCTTAATGAGTGAGCTAACTCA 831
Db 97 ACATACGAGCCGGAAGCATAAAGTGAAGCTGGGTGCTTAATGAGTGAGCTAACTCA 156

QY 832 CATTAATTCGTTGCGTCACTGCCCGCTTTCCAGTCGGGAACCTGTCGTCGAGCTGC 891
Db 157 CATTAATTCGTTGCGTCACTGCCCGCTTTCCAGTCGGGAACCTGTCGTCGAGCTGC 216

QY 892 ATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTCGTTATTCGGCGCTCTTCCGCTT 951
Db 217 ATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTCGTTATTCGGCGCTCTTCCGCTT 276

QY 952 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
Db 277 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 326

RESULT 13

ABX99895
ID ABX99895 standard; cDNA; 584 BP.

XX AC
XX AC
XX ABX99895;

XX 22-MAY-2003 (first entry)

XX Lung cancer therapy and diagnosis associated cDNA #622.

XX Lung cancer; cytostatic; vaccine; gene therapy; cancer;
XX gene; ss.

XX Homo sapiens.

XX US2002172952-A1.

XX 21-NOV-2002.

XX 10-JUL-2001; 2001US-0902941.

XX PF
XX PD
XX DE

PR 30-JUN-1999; 99US-0346492.
PR 15-OCT-1999; 99US-0419356.
PR 17-DEC-1999; 99US-0468867.
PR 30-DEC-1999; 99US-0476300.
PR 06-MAR-2000; 2000US-0519642.
PR 22-MAR-2000; 2000US-0531077.
PR 10-APR-2000; 2000US-0546259.
PR 27-APR-2000; 2000US-0560406.
PR 05-JUN-2000; 2000US-0589184.
PR 11-JUL-2000; 2000US-0614124.
PR 29-AUG-2000; 2000US-0651563.
PR 08-SEP-2000; 2000US-0658824.
PR 26-SEP-2000; 2000US-0671325.
PR 06-OCT-2000; 2000US-0677419.
PR 30-OCT-2000; 2000US-0702705.
PR 13-DEC-2000; 2000US-0736457.
PR 03-MAY-2001; 2001US-0849626.
XX (CORI-) CORIXA CORP.

XX Henderson RA, Wang T, Watanabe Y, Johnson JC, Retter MW, Durham M;
PI Carter D, Fanger GR, Vedwick TS, Bangur CS, McNabb A;

XX WPI; 2003-328427/31.

XX New polynucleotide, useful for preparing a composition for treating or
PT inhibiting development of cancer, e.g. lung cancer -

PS Example 1; SEQ ID NO 639; 82pp; English.

XX The invention describes an isolated polynucleotide comprising one of 32
CC sequences, complement or degenerate variants of them. The polynucleotide
CC is useful for preparing a composition e.g. a vaccine or for gene therapy,
CC for treating or inhibiting development of cancer, e.g. lung cancer.
CC This sequence represents a polynucleotide associated with the
CC compositions and methods for the therapy and diagnosis of lung cancer.

XX Sequence 584 BP; 142 A; 159 C; 153 G; 130 T; 0 other;

Query Match 29.0%; Score 290; DB 25; Length 584;
Best Local Similarity 100.0%; Pred. No. 1.1e-80;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGTCATAGCTGTTTCTGTCGTAATTTGTTATCCGCTCACAAATCCACACA 771
Db 37 CGTAATCATGTCATAGCTGTTTCTGTCGTAATTTGTTATCCGCTCACAAATCCACACA 96

QY 772 ACATACGAGCCGGAAGCATAAAGTGAAGCTGGGTGCTTAATGAGTGAGCTAACTCA 831
Db 97 ACATACGAGCCGGAAGCATAAAGTGAAGCTGGGTGCTTAATGAGTGAGCTAACTCA 156

QY 832 CATTAATTCGTTGCGTCACTGCCCGCTTTCCAGTCGGGAACCTGTCGTCGAGCTGC 891
Db 157 CATTAATTCGTTGCGTCACTGCCCGCTTTCCAGTCGGGAACCTGTCGTCGAGCTGC 216

QY 892 ATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTCGTTATTCGGCGCTCTTCCGCTT 951
Db 217 ATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTCGTTATTCGGCGCTCTTCCGCTT 276

QY 952 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
Db 277 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 326

RESULT 14

AAA62566
ID AAA62566 standard; cDNA; 598 BP.

XX AC
XX AAA62566;

XX 14-NOV-2000 (first entry)

XX Human breast cancer cell cDNA sequence #382.

XX Human; arbitrary primer; cDNA synthesis; contig sequence construction;
KW open reading frame; ORF; low stringency; cDNA sequencing; ss.
XX Homo sapiens.
OS WO200031299-A2.
PN 02-JUN-2000.
XX 19-NOV-1999; 99WO-US27430.
XX 20-NOV-1998; 98US-0196716.
XX (LUDW-) LUDWIG INST CANCER RES.
XX Simpson AJG, Dias Neto E, Brentani RR;
PI WPI; 2000-400100/34.
XX
XX Method for determining open reading frames of the genome of an organism
PT using primers at low stringency conditions, useful in the construction
PT of contigs or constructs of sequenced nucleic acid molecules -
XX
XX Example 7; Page 112; 113pp; English.
XX
XX The present sequence is a cDNA sequence obtained using a method for
CC determining open reading frames (ORFs) of the genome of an
CC organism. An aliquot of mRNA from human breast cancer cells was mixed
CC with a single, arbitrary primer, Moloney murine leukaemia virus reverse
CC transcriptase, reverse transcriptase buffer and dNTPs. The mixture was
CC incubated under low stringency conditions to yield single stranded
CC cDNA. The same primer was then used to amplify the cDNA by PCR. Rather
CC than providing nucleotide sequence information from the non-coding
CC termini of nucleic acid molecules, the method provides information on
CC the more interesting and relevant internal portions, such as ORFs. The
CC method also permits the construction of contigs of sequenced nucleic
CC acid molecules.
XX
XX Sequence 598 BP; 136 A; 170 C; 155 G; 132 T; 5 other;
SQ
Query Match 29.0%; Score 290; DB 21; Length 598;
Best Local Similarity 100.0%; Pred. No. 1.1e-80;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 712 CGTAATCATGTCTAGCTGTTCTCTGTGGAATTCTTATCCGCTCACAAATTCACACA 771
Db 10 CGTAATCATGTCTAGCTGTTCTCTGTGGAATTCTTATCCGCTCACAAATTCACACA 69
QY 772 ACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGTGCTTAATGAGTGAGTAACTCA 831
Db 70 ACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGTGCTTAATGAGTGAGTAACTCA 129
QY 832 CATTAATTCGGTTCGCTCACTGCCCGCTTCCAGTCGGGAAACCTGCTCGGAGCTGC 891
Db 130 CATTAATTCGGTTCGCTCACTGCCCGCTTCCAGTCGGGAAACCTGCTCGGAGCTGC 189
QY 892 ATTAATGAATCGGCAACGCGCGGGAGAGCGGTTTCGCTATTGGCGCTCTTCGCTT 951
Db 190 ATTAATGAATCGGCAACGCGCGGGAGAGCGGTTTCGCTATTGGCGCTCTTCGCTT 249
QY 952 CTGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
Db 250 CTGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 299
RESULT 15
ABT12681
ID ABT12681 standard; DNA; 598 BP.
XX
AC ABT12681;
XX
DT 02-JAN-2003 (first entry)

XX Orestes sequencing method related DNA SEQ ID No 387.
DE
XX Open reading frame; ORF; low stringency; sequencing; ORESTES; bronchitis;
KW identification; internal region; population study; heart disease;
KW Alzheimer's disease; autoimmune disease; congenital disease; human; ds.
XX
OS Homo sapiens.
XX WO200274994-A2.
PN 26-SEP-2002.
XX 01-NOV-2001; 2001WO-US46665.
XX 07-NOV-2000; 2000US-246313P.
XX (LUDW-) LUDWIG INST CANCER RES.
XX Simpson AJG, Dias-neto E, Brentani RR;
PI WPI; 2002-750567/81.
XX
XX Determining ORFs of the genome by contacting mRNA with a single
PT oligonucleotide primer, useful for population studies that identify
PT genes associated with diseases such as cardiovascular, autoimmune and
PT congenital diseases -
XX
XX Example 7; Page 161-162; 166pp; English.
XX
XX The invention relates to a novel method for determining open reading
CC frames (ORFs) of the genome of an organism comprising contacting
CC messenger RNA from a cell of the organism with a single, oligonucleotide
CC primer at low stringency. The method is an improved ORESTES sequencing
CC method. The methods of the present invention are useful for sequencing
CC of, essentially, the entire coding component of an organism,
CC identification of previously unknown nucleic acid molecules, assigning
CC previously identified nucleotide sequences to internal regions of genes
CC and population studies that identify genes associated with diseases such
CC as heart disease, bronchitis, Alzheimer's disease, autoimmune diseases
CC and congenital diseases. This polynucleotide represents a DNA sequence
CC that is used in the novel method of the invention.
XX
XX Sequence 598 BP; 136 A; 170 C; 155 G; 132 T; 5 other;
SQ
Query Match 29.0%; Score 290; DB 24; Length 598;
Best Local Similarity 100.0%; Pred. No. 1.1e-80;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 712 CGTAATCATGTCTAGCTGTTCTCTGTGGAATTCTTATCCGCTCACAAATTCACACA 771
Db 10 CGTAATCATGTCTAGCTGTTCTCTGTGGAATTCTTATCCGCTCACAAATTCACACA 69
QY 772 ACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGTGCTTAATGAGTGAGTAACTCA 831
Db 70 ACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGTGCTTAATGAGTGAGTAACTCA 129
QY 832 CATTAATTCGGTTCGCTCACTGCCCGCTTCCAGTCGGGAAACCTGCTCGGAGCTGC 891
Db 130 CATTAATTCGGTTCGCTCACTGCCCGCTTCCAGTCGGGAAACCTGCTCGGAGCTGC 189
QY 892 ATTAATGAATCGGCAACGCGCGGGAGAGCGGTTTCGCTATTGGCGCTCTTCGCTT 951
Db 190 ATTAATGAATCGGCAACGCGCGGGAGAGCGGTTTCGCTATTGGCGCTCTTCGCTT 249
QY 952 CTGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
Db 250 CTGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 299
Search completed; February 15, 2004, 22:40:26
Job time : 327.065 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 10:22:55 ; Search time 73.6147 Seconds
(without alignments)
6001.851 Million cell updates/sec

Title: US-09-921-143-36_COPY_2000_3000

Perfect score: 1001

Sequence: 1 tctcagggttaaccagtgggg.....ttcggctggcgagcggtta 1001

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 569978 seqs, 220691566 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/2/ina/5A COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/5B COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/5A COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/5B COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/PTCUS COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	307	30.7	5552	3	US-08-155-888-1
2	290	29.0	450	4	US-09-486-336A-2
3	290	29.0	584	4	US-09-702-705-639
4	290	29.0	584	4	US-09-736-457-639
5	290	29.0	701	2	US-08-706-702-24
6	290	29.0	701	3	US-08-706-706-24
7	290	29.0	701	4	US-09-238-471-24
8	290	29.0	713	2	US-08-706-702-26
9	290	29.0	713	3	US-08-706-706-26
10	290	29.0	713	4	US-09-238-471-26
11	290	29.0	716	3	US-08-998-416-55
12	290	29.0	717	2	US-08-706-702-25
13	290	29.0	717	3	US-08-706-706-25
14	290	29.0	717	4	US-09-238-471-25
15	290	29.0	738	1	US-08-726-462B-3
16	290	29.0	738	2	US-09-046-203-3
17	290	29.0	738	3	US-09-272-104-3
18	290	29.0	738	4	US-09-272-097-3
19	290	29.0	752	4	US-08-976-259-108
20	290	29.0	809	4	US-09-328-475C-259
21	290	29.0	848	4	US-09-328-475C-195
22	290	29.0	1024	4	US-09-328-475C-70
23	290	29.0	1738	1	US-07-971-160-9
24	290	29.0	1738	1	US-08-336-241-9
25	290	29.0	1738	2	US-08-465-273-9
26	290	29.0	1738	2	US-09-119-024-9
27	290	29.0	1738	2	US-08-417-226-9

Sequence 9, Appli
Sequence 9, Appli
Sequence 7, Appli
Sequence 6, Appli
Sequence 11, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 1, Appli
Sequence 32, Appli
Sequence 4, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 21, Appli
Sequence 33, Appli
Sequence 16, Appli

ALIGNMENTS

RESULT 1

US-08-155-888-1
; Sequence 1, Application US/08155888
; Patent No. 6066223
; GENERAL INFORMATION:
; APPLICANT: Hoffman, Stephen L.
; APPLICANT: Hedstrom, Richard C.
; APPLICANT: Sedegah, Martha
; TITLE OF INVENTION: POLYNUCLEOTIDE VACCINE PROTECTIVE
; AGAINST MALARIA, METHODS OF PROTECTION AND VECTOR FOR
; DELIVERING POLYNUCLEOTIDE VACCINES
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Naval Medical Res. & Dev. Cmd.
; STREET: Bldg. 1, T-12 8901 Wisconsin Ave.
; CITY: Bethesda
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20889-5606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/155,888
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Spevack, A. David
; REGISTRATION NUMBER: 24,743
; REFERENCE/DOCKET NUMBER: N.C. 75,851
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 295-6759
; TELEFAX: (202) 295-1022
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5552 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: Promoter
; LOCATION: 1..755
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /function= "promoter"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /label= CMV-IE

QY 832 CATTAATGTCGTCGCTCACTCCCGCTTCCAGTCGGGAAACCTGTCGTCAGCTGC 891
Db 167 CATTAATGTCGTCGCTCACTCCCGCTTCCAGTCGGGAAACCTGTCGTCAGCTGC 226
QY 892 ATTAATGAATCGGCCCAACGCGCGGAGAGCGGTTTGGCTATTGGCGCTCTTCCGCTT 951
Db 227 ATTAATGAATCGGCCCAACGCGCGGAGAGCGGTTTGGCTATTGGCGCTCTTCCGCTT 286
QY 952 CCTCGCTCACTGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
Db 287 CCTCGCTCACTGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 336

RESULT 6

US-08-706-706-24
; Sequence 24, Application US/08706706
; Patent No. 6015668
; GENERAL INFORMATION:
; APPLICANT: Hughes, A. John
; TITLE OF INVENTION: Cloned DNA Polymerases from Thermotoga and
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
; STREET: 1100 New York Ave., N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/706,706
; FILING DATE: 06-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/689,807
; FILING DATE: 14-AUG-1996
; PRIOR APPLICATION NUMBER: US 08/537,400
; FILING DATE: 02-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/370,190
; FILING DATE: 01-JAN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/316,423
; FILING DATE: 30-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/576,759
; FILING DATE: 21-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/537,397
; FILING DATE: 02-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/525,057
; FILING DATE: 08-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.2800005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 701 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: DNA (genomic)

US-08-706-706-24

Query Match 29.0%; Score 290; DB 3; Length 701;
Best Local Similarity 100.0%; Pred. No. 1.5e-105;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 712 CGTAATCATGTCATAGCTGTTTCTCGTGTGAATTTGTAATCCGCTCACAAATTCACACA 771
Db 47 CGTAATCATGTCATAGCTGTTTCTCGTGTGAATTTGTAATCCGCTCACAAATTCACACA 106
QY 772 ACATACGAGCCGGAAGCATAAAGTGTAAAGCTCGGGTGCCTAATAGTAGAGCTAACTCA 831
Db 107 ACATACGAGCCGGAAGCATAAAGTGTAAAGCTCGGGTGCCTAATAGTAGAGCTAACTCA 166
QY 832 CATTAATGTCGTCGCTCACTCCCGCTTCCAGTCGGGAAACCTGTCGTCAGCTGC 891
Db 167 CATTAATGTCGTCGCTCACTCCCGCTTCCAGTCGGGAAACCTGTCGTCAGCTGC 226
QY 892 ATTAATGAATCGGCCCAACGCGCGGAGAGCGGTTTGGCTATTGGCGCTCTTCCGCTT 951
Db 227 ATTAATGAATCGGCCCAACGCGCGGAGAGCGGTTTGGCTATTGGCGCTCTTCCGCTT 286
QY 952 CCTCGCTCACTGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
Db 287 CCTCGCTCACTGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 336

RESULT 7

US-09-238-471-24
; Sequence 24, Application US/09238471
; Patent No. 6506560
; GENERAL INFORMATION:
; APPLICANT: Hughes, A. John
; TITLE OF INVENTION: Cloned DNA Polymerases from Thermotoga and
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
; STREET: 1100 New York Ave., N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/09/238,471
; FILING DATE: 28-Jan-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/706,706
; FILING DATE: 06-SEP-1996
; APPLICATION NUMBER: US 08/689,807
; FILING DATE: 14-AUG-1996
; APPLICATION NUMBER: US 08/537,400
; FILING DATE: 02-OCT-1995
; APPLICATION NUMBER: US 08/370,190
; FILING DATE: 01-JAN-1995
; APPLICATION NUMBER: US 08/316,423
; FILING DATE: 30-SEP-1994
; APPLICATION NUMBER: US 08/576,759
; FILING DATE: 21-DEC-1995
; APPLICATION NUMBER: US 08/537,397
; FILING DATE: 02-OCT-1995
; APPLICATION NUMBER: US 08/525,057
; FILING DATE: 08-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.2800005

```

; APPLICATION NUMBER: US 08/537,373
; FILING DATE: 02-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/525,057
; FILING DATE: 08-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.2800006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 26:
; .SEQUENCE CHARACTERISTICS:
; LENGTH: 713 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLSCULE TYPE: DNA (genomic)
; US-08-706-702-26

      29.0%; Score 290; DB 2; Length 713;
Best Local Similarity 100.0%; Pred. No. 1.4e-105;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      712 CGTAATCATGGTCATAGCTGTTTCCTGTGTGAAGTATGATCCCGTCACAATTCACACA 771
Db      81 CGTAATCATGGTCATAGCTGTTTCCTGTGTGAAGTATGATCCCGTCACAATTCACACA 140
QY      772 ACATACGAGCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTTAATGAGTGAGCTAACTCA 831
Db      141 ACATACGAGCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTTAATGAGTGAGCTAACTCA 200
QY      832 CATTAATGGGTGGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGGCCAGCTGC 891
Db      201 CATTAATGGGTGGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGGCCAGCTGC 260
QY      892 ATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTCGTATATGGGCGCTTTCCGCTT 951
Db      261 ATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTCGTATATGGGCGCTTTCCGCTT 320
QY      952 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 1001
Db      321 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 370

RESULT 9
US-08-706-706-26
; Sequence 26, Application US/08706706
; Patent No. 6015668
; GENERAL INFORMATION:
; APPLICANT: Hughes, A. John
; APPLICANT: Chatterjee, Deb K.
; TITLE OF INVENTION: Cloned DNA Polymerases from Thermotoga and
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
; STREET: 1100 New York Ave., N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/706,706
; FILING DATE: 06-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/689,807

```


Sequence 55, Application US/08999416
Patent No. 6239264
GENERAL INFORMATION:
APPLICANT: Philippsen, Peter
APPLICANT: Pohlmann, Sabine
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtel, Philipp
APPLICANT: Reibschung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPHII
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 716 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORGANISM: PAG1015UP
US-08-998-416-55

Query Match 29.0%; Score 290; DB 3; Length 716;
Best Local Similarity 100.0%; Pred. No. 1.4e-105;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 712 CGTAATCATGGTCATAGCTGTTTCCTGTGTGAATTTATCCGCTCACAATTCACACA 771
DB 206 CGTAATCATGGTCATAGCTGTTTCCTGTGTGAATTTATCCGCTCACAATTCACACA 265
QY 772 ACATACAGCGGAGCATAAAGTGTAAAGCTGGGGTGCTTAATGAGTGAGCTAACTCA 831
DB 266 ACATACAGCGGAGCATAAAGTGTAAAGCTGGGGTGCTTAATGAGTGAGCTAACTCA 325
QY 832 CATTAATTGCGTTCAGTCCCGCTTTCCTGTCGGGAAACCTGTGTCGCGAGCTGC 891
DB 326 CATTAATTGCGTTCAGTCCCGCTTTCCTGTCGGGAAACCTGTGTCGCGAGCTGC 385
QY 892 ATTAATGATTCGGCAACCGCGGGAGAGCGGTTTCGTTATGGGGCTTTCGCTT 951
DB 386 ATTAATGATTCGGCAACCGCGGGAGAGCGGTTTCGTTATGGGGCTTTCGCTT 445
QY 952 CCTCGCTCACTGACTCGCTCGCTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTT 1001
DB 446 CCTCGCTCACTGACTCGCTCGCTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTT 495

RESULT 12
US-08-706-702-25
Sequence 25, Application US/08706702
Patent No. 5948614
GENERAL INFORMATION:
APPLICANT: Chatterjee, Deb K.
TITLE OF INVENTION: Cloned DNA Polymerases from Thermotoga
TITLE OF INVENTION: maritima and Mutants Thereof
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/706,702
FILING DATE: 06-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/689,807
FILING DATE: 14-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/537,400
FILING DATE: 02-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/576,759
FILING DATE: 21-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/537,397
FILING DATE: 02-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/525,057
FILING DATE: 08-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.2800006
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 717 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: DNA (genomic)
US-08-706-702-25

Query Match 29.0%; Score 290; DB 2; Length 717;
Best Local Similarity 100.0%; Pred. No. 1.4e-105;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 712 CGTAATCATGGTCATAGCTGTTTCCTGTGTGAATTTATCCGCTCACAATTCACACA 771
DB 85 CGTAATCATGGTCATAGCTGTTTCCTGTGTGAATTTATCCGCTCACAATTCACACA 144
QY 772 ACATACAGCGGAGCATAAAGTGTAAAGCTGGGGTGCTTAATGAGTGAGCTAACTCA 831
DB 145 ACATACAGCGGAGCATAAAGTGTAAAGCTGGGGTGCTTAATGAGTGAGCTAACTCA 204
QY 832 CATTAATTGCGTTCGCTCACTGCCCGCTTTCAGTTCGGGAAACCTGTGTCGCGAGCTGC 891
DB 205 CATTAATTGCGTTCGCTCACTGCCCGCTTTCAGTTCGGGAAACCTGTGTCGCGAGCTGC 264

QY 892 ATTAATGAATCGCCCAACCGCGGAGAGCGGTTTCCGTTATGGCGCTCTTCGGCTT 951
Db 265 ATTAATGAATCGCCCAACCGCGGAGAGCGGTTTCCGTTATGGCGCTCTTCGGCTT 324
QY 952 CTTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1001
Db 325 CTTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 374

RESULT 13

US-08-706-706-25
; Sequence 25, Application US/08706706
; Patent No. 6015668
; GENERAL INFORMATION:
; APPLICANT: Hughes, A. John
; TITLE OF INVENTION: Cloned DNA Polymerases from Thermotoga and
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
; STREET: 1100 New York Ave., N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/706,706
FILING DATE: 06-SEP-1996
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/689,807
FILING DATE: 14-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/537,400
FILING DATE: 02-OCT-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/370,190
FILING DATE: 01-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/316,423
FILING DATE: 30-SEP-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/576,759
FILING DATE: 21-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/537,397
FILING DATE: 02-OCT-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/525,057
FILING DATE: 08-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Esmord, Robert W.

REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.2800005
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2540
TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 717 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both

MOLECULAR TYPE: DNA (genomic)
US-08-706-706-25

Query Match

29.0%; Score 290; DB 3; Length 717;

Best Local Similarity 100.0%; Fred. No. 1.4e-105;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 712 CGTAATCATGCTCATAGCTGTTTCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 771
Db 85 CGTAATCATGCTCATAGCTGTTTCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 144
QY 772 ACATACAGCGCGAAGCATATAAGCTGTAAAGCTGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 831
Db 145 ACATACAGCGCGAAGCATATAAGCTGTAAAGCTGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 204
QY 832 CATTAATTGGCTTGGCTCACTGCCCGCTTCCAGTCCGGAAACCTGTCGTGCGCAGCTGC 891
Db 205 CATTAATTGGCTTGGCTCACTGCCCGCTTCCAGTCCGGAAACCTGTCGTGCGCAGCTGC 264
QY 892 ATTAATGAATCGCCCAACCGCGGAGAGCGGTTTCCGTTATGGCGCTCTTCGGCTT 951
Db 265 ATTAATGAATCGCCCAACCGCGGAGAGCGGTTTCCGTTATGGCGCTCTTCGGCTT 324
QY 952 CTTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1001
Db 325 CTTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 374

RESULT 14

US-09-238-471-25
; Sequence 25, Application US/09238471
; Patent No. 6505560
; GENERAL INFORMATION:
; APPLICANT: Hughes, A. John

TITLE OF INVENTION: Cloned DNA Polymerases from Thermotoga and
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/238,471
FILING DATE: 28-Jan-1999
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/706,706
FILING DATE: 06-SEP-1996
APPLICATION NUMBER: US 08/689,807
FILING DATE: 14-AUG-1996
APPLICATION NUMBER: US 08/537,400
FILING DATE: 02-OCT-1995
APPLICATION NUMBER: US 08/370,190
FILING DATE: 01-JAN-1995
APPLICATION NUMBER: US 08/316,423
FILING DATE: 30-SEP-1994
APPLICATION NUMBER: US 08/576,759
FILING DATE: 21-DEC-1995
APPLICATION NUMBER: US 08/537,397
FILING DATE: 02-OCT-1995
APPLICATION NUMBER: US 08/525,057
FILING DATE: 08-SEP-1995

ATTORNEY/AGENT INFORMATION:
NAME: Esmord, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.2800005
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540

;; INFORMATION FOR SEQ ID NO: 25:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 717 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: both
;; TOPOLOGY: both
;; MOLECULE TYPE: DNA (genomic)
;; SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-238-471-25

Query Match 29.0%; Score 290; DB 4; Length 717;
Best Local Similarity 100.0%; Pred. No. 1.4e-105;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGGTCATAGCTGTTTCCTGTTGTAATTTGTTATCCGCTCACAATTCACACA 771
DB 85 CGTAATCATGGTCATAGCTGTTTCCTGTTGTAATTTGTTATCCGCTCACAATTCACACA 144
QY 772 ACATAGAGCCGGAAGCATTAAGTGAAGCTGGGTGCTTAATGAGTGAGCTTAATCA 831
DB 145 ACATAGAGCCGGAAGCATTAAGTGAAGCTGGGTGCTTAATGAGTGAGCTTAATCA 204
QY 832 CATTAATTCGTTGCGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGC 891
DB 205 CATTAATTCGTTGCGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGC 264
QY 892 ATTAATGAATCGGCAACCGCGGGAGAGCGGTTTGGCTATTGGCGCTCTTCGCTT 951
DB 265 ATTAATGAATCGGCAACCGCGGGAGAGCGGTTTGGCTATTGGCGCTCTTCGCTT 324
QY 952 CTCGCTCACTGACTCGTGGCTCGGTCGTTCCGCTGCGGCGAGCGGTA 1001
DB 325 CTCGCTCACTGACTCGTGGCTCGGTCGTTCCGCTGCGGCGAGCGGTA 374

RESULT 15
US-08-726-462B-3
; Sequence 3, Application US/08726462B
; Patent No. 5800996
; GENERAL INFORMATION:
; APPLICANT: Perkin-Elmer Corporation, Applied Biosystems
; APPLICANT: Division
; TITLE OF INVENTION: ENERGY TRANSFER DYES WITH ENHANCED
; TITLE OF INVENTION: FLUORESCENCE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David J. Weitz, Wilson Sonsini Goodrich
; ADDRESSEE: & Rosati
; STREET: 650 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1050
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: Microsoft Windows 3.1/DOS 5.0
; SOFTWARE: Wordperfect for Windows 6.0,
; SOFTWARE: ASCII (DOS) TEXT format
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,462B
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/642,330
; FILING DATE: May 3, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/672,196
; FILING DATE: June 27, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: David J. Weitz
; REGISTRATION NUMBER: 38,362
; REFERENCE/DOCKET NUMBER: PELM4304

;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 493-9300
;; TELEFAX: (415) 493-6811
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 738 nucleotides
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-08-726-462B-3

Query Match 29.0%; Score 290; DB 1; Length 738;
Best Local Similarity 100.0%; Pred. No. 1.4e-105;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGGTCATAGCTGTTTCCTGTTGTAATTTGTTATCCGCTCACAATTCACACA 771
DB 109 CGTAATCATGGTCATAGCTGTTTCCTGTTGTAATTTGTTATCCGCTCACAATTCACACA 168
QY 772 ACATAGAGCCGGAAGCATTAAGTGAAGCTGGGTGCTTAATGAGTGAGCTTAATCA 831
DB 169 ACATAGAGCCGGAAGCATTAAGTGAAGCTGGGTGCTTAATGAGTGAGCTTAATCA 228
QY 832 CATTAATTCGTTGCGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGC 891
DB 229 CATTAATTCGTTGCGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGC 288
QY 892 ATTAATGAATCGGCAACCGCGGGAGAGCGGTTTGGCTATTGGCGCTCTTCGCTT 951
DB 289 ATTAATGAATCGGCAACCGCGGGAGAGCGGTTTGGCTATTGGCGCTCTTCGCTT 348
QY 952 CTCGCTCACTGACTCGTGGCTCGGTCGTTCCGCTGCGGCGAGCGGTA 1001
DB 349 CTCGCTCACTGACTCGTGGCTCGGTCGTTCCGCTGCGGCGAGCGGTA 398

Search completed: February 15, 2004, 19:10:14
Job time : 75.6147 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 16:21:35 ; Search time 420.484 Seconds
(without alignments)
8769.219 Million cell updates/sec

Title: US-09-921-143-36_COPY_2000_3000

Perfect score: 1001

Sequence: 1 tctcaggttaacagctggggg.....ttcggtcgccgagcggtta 1001

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 2449703 seqs, 1841816367 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubna/US09_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubna/US10_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubna/US10_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubna/US10_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq.*
17: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1001	100.0	5283	13	US-09-921-143-36
2	291	29.1	9632	12	US-10-394-511-1
3	290	29.0	380	10	US-09-796-652-3110
4	290	29.0	380	12	US-10-057-475B-3110
5	290	29.0	380	12	US-10-154-864B-3110
6	290	29.0	380	15	US-10-040-862-3110
7	290	29.0	413	13	US-10-270-839-103
8	290	29.0	557	10	US-09-796-652-3746
9	290	29.0	557	12	US-10-057-475B-3746
10	290	29.0	557	12	US-10-154-864B-3746
11	290	29.0	557	15	US-10-040-862-3746
12	290	29.0	584	10	US-09-736-457-639
13	290	29.0	584	10	US-09-902-941-639
14	290	29.0	584	10	US-09-849-626-639
15	290	29.0	584	11	US-09-476-300-639

16	290	29.0	584	13	US-10-113-872-639	Sequence 639, App
17	290	29.0	584	15	US-10-017-754-639	Sequence 639, App
18	290	29.0	605	13	US-09-814-353-18505	Sequence 18505, A
19	290	29.0	686	10	US-09-764-868-1382	Sequence 1382, Ap
20	290	29.0	686	11	US-09-764-891-8687	Sequence 8687, Ap
21	290	29.0	686	11	US-09-764-891-9713	Sequence 9713, Ap
22	290	29.0	701	11	US-09-229-173-24	Sequence 24, Appl
23	290	29.0	701	11	US-10-285-696-24	Sequence 24, Appl
24	290	29.0	713	11	US-09-229-173-26	Sequence 26, Appl
25	290	29.0	713	13	US-10-285-696-26	Sequence 26, Appl
26	290	29.0	717	11	US-09-229-173-25	Sequence 25, Appl
27	290	29.0	717	13	US-10-285-696-25	Sequence 25, Appl
28	290	29.0	752	9	US-09-956-004-108	Sequence 108, App
29	290	29.0	803	13	US-10-257-826A-17	Sequence 17, Appl
30	290	29.0	808	15	US-10-198-846-6873	Sequence 6873, Ap
31	290	29.0	809	14	US-10-202-193-259	Sequence 259, App
32	290	29.0	848	14	US-10-202-193-195	Sequence 195, App
33	290	29.0	1024	14	US-10-202-193-70	Sequence 70, Appl
34	290	29.0	1091	12	US-09-764-875-184	Sequence 184, App
35	290	29.0	1092	10	US-09-764-868-232	Sequence 232, App
36	290	29.0	1092	11	US-09-764-891-1853	Sequence 1853, Ap
37	290	29.0	1092	12	US-09-764-875-482	Sequence 482, App
38	290	29.0	1355	9	US-09-764-870-182	Sequence 182, App
39	290	29.0	1355	9	US-09-764-853-404	Sequence 404, App
40	290	29.0	1355	12	US-10-158-057-165	Sequence 165, App
41	290	29.0	1355	15	US-10-125-540-182	Sequence 182, App
42	290	29.0	1355	15	US-10-103-313-273	Sequence 273, App
43	290	29.0	1738	11	US-09-836-169-9	Sequence 9, Appl
44	290	29.0	1738	13	US-10-152-994-9	Sequence 9, Appl
45	290	29.0	2074	15	US-10-106-698-350	Sequence 350, App

ALIGNMENTS

RESULT 1
US-09-921-143-36
; Sequence 36, Application US/09921143
; Publication No: US20030215921A1
; GENERAL INFORMATION:
; APPLICANT: Coleman, Timothy
; TITLE OF INVENTION: Vascular Endothelial Growth Factor-2
; FILE REFERENCE: PFI12P6
; CURRENT APPLICATION NUMBER: US/09/921.143
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/223,276
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 36
; LENGTH: 5283
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-921-143-36

Query Match	100.0%	Score 1001;	DB 13;	Length 5283;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1001;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	TCTCAGCTAACAGTGGGGGCGACATTCCTGCGGAGCTAGACATATGTAACATGCT	60	
Db	2000	TCTCAGCTAACAGTGGGGGCGACATTCCTGCGGAGCTAGACATATGTAACATGCT	2059	
Qy	61	AGTGCAGGAAGAGTGAAGATCCCTTAACTCTCTAGTGTGGTGGGCTAG	120	
Db	2060	AGTGCAGGAAGAGTGAAGATCCCTTAACTCTCTAGTGTGGTGGGCTAG	2119	
Qy	121	GCCCCAGGATAGTACCTATTGGGGAGCCCATAGACCTGCACTGACCTAGGATGCT	180	
Db	2120	GCCCCAGGATAGTACCTATTGGGGAGCCCATAGACCTGCACTGACCTAGGATGCT	2179	
Qy	181	AACAGATGTAGTGGTTTGGAGGCCCATATGTCCTATTCAGACCTGACTGCTTCAC	240	

Db 2180 AACAGGATGTAGTGTGGAGGCCAATATGTCATTCATGACCAAGTGTGTCTCAC 2235
Qy 241 AGCCATCAACCCCTTGTCTCTGTGCTGACTTATAGCAGGGGATAAAGTGAGAGAAAGCCCTG 300
Db 2240 AGCCATCAACCCCTTGTCTCTGTGCTGACTTATAGCAGGGGATAAAGTGAGAGAAAGCCCTG 2299
Qy 301 GGCTAAATCAGGGGCTGCTCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
Db 2300 GGCTAAATCAGGGGCTGCTCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2359
Qy 361 GTGCTGTGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 420
Db 2360 GTGCTGTGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 2419
Qy 421 TGGGTGAGAGCCCGGGGAGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 480
Db 2420 TGGGTGAGAGCCCGGGGAGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 2479
Qy 481 AGCGGGCATCGTGTGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 540
Db 2480 AGCGGGCATCGTGTGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 2539
Qy 541 ACTGCAACTAGGCCCCACCACTACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 2540 ACTGCAACTAGGCCCCACCACTACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2599
Qy 601 GAGCACTACAAGTTGTGTGTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 2600 GAGCACTACAAGTTGTGTGTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2659
Qy 661 GTGGGGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 720
Db 2660 GTGGGGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 2719
Qy 721 GGTCACTAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 780
Db 2720 GGTCACTAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 2779
Qy 781 CCGCAAGCATAAAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 840
Db 2780 CCGCAAGCATAAAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 2839
Qy 841 CGTTGCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 2840 CGTTGCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2899
Qy 901 TCGGCCAACCGCGGGGAGAGCGGTTTCCAGTTCGGGAAACCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 2900 TCGGCCAACCGCGGGGAGAGCGGTTTCCAGTTCGGGAAACCTGCTGCTGCTGCTGCTGCTGCTGCT 2959
Qy 961 CTGACTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1001
Db 2960 CTGACTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000

RESULT 2

US-10-394-511-1
; Sequence 1, Application US/10394511
; Publication No. US20040009551A1
; GENERAL INFORMATION:
; APPLICANT: Borsford, Thor
; TITLE OF INVENTION: Ricin-like Toxin Variants for Treatment of Cancer,
; FILE OF INVENTION: Viral or Parasitic Infections
; FILE REFERENCE: 10447-27
; CURRENT APPLICATION NUMBER: US/10/394,511
; CURRENT FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: US 09/403,752
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: US 60/063,715
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: US 60/045,148
; PRIOR FILING DATE: 1997-10-30
; NUMBER OF SEQ ID NOS: 142

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 9632

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Baculovirus transfer vector pVLL393

US-10-394-511-1

Query Match

Best Local Similarity 100.0%; Pred. No. 3.4e-144;

Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 711 ACCTAATCATGTGCTAGTCTTTCTCTGTGTGAATTTGTTATCGCTCACAATTTCCACAC 770

Db 7002 ACCTAATCATGTGCTAGTCTTTCTCTGTGTGAATTTGTTATCGCTCACAATTTCCACAC 7061

Qy 771 AACATACGAGCCGAGCATAAAGTGTAAAGCTTGGGTGCTTAATGAGTGAATCTC 830

Db 7062 AACATACGAGCCGAGCATAAAGTGTAAAGCTTGGGTGCTTAATGAGTGAATCTC 7121

Qy 831 ACATTAATTCGCTTGGCTCTCACTGCCGCTTTCCAGTCGGGAAACCTGCTGCCAGCTG 890

Db 7122 ACATTAATTCGCTTGGCTCTCACTGCCGCTTTCCAGTCGGGAAACCTGCTGCCAGCTG 7181

Qy 891 CATTAATGAATCGCCAAACGCGGGGAGAGCGGTTTGGTATTGGGCGCTCTTCCGCT 950

Db 7182 CATTAATGAATCGCCAAACGCGGGGAGAGCGGTTTGGTATTGGGCGCTCTTCCGCT 7241

Qy 951 TCTCTCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1001

Db 7242 TCTCTCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7292

RESULT 3

US-09-796-692-3110

; Sequence 3110, Application US/09796692

; Publication No. US20020198362A1

; GENERAL INFORMATION:

; APPLICANT: Gaiger, Alexander

; APPLICANT: Mannion, Jane

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER

; FILE OF INVENTION: HEMATOLOGICAL MALIGNANCIES

; FILE REFERENCE: 2077.001200

; CURRENT APPLICATION NUMBER: US/09/796,692

; CURRENT FILING DATE: 2001-03-01

; PRIOR APPLICATION NUMBER: 60/186,126

; PRIOR FILING DATE: 2000-03-01

; PRIOR APPLICATION NUMBER: 60/190,479

; PRIOR FILING DATE: 2000-03-17

; PRIOR APPLICATION NUMBER: 60/200,545

; PRIOR FILING DATE: 2000-04-27

; PRIOR APPLICATION NUMBER: 60/200,303

; PRIOR FILING DATE: 2000-04-28

; PRIOR APPLICATION NUMBER: 60/200,779

; PRIOR FILING DATE: 2000-04-28

; PRIOR APPLICATION NUMBER: 60/200,999

; PRIOR FILING DATE: 2000-05-01

; PRIOR APPLICATION NUMBER: 60/202,084

; PRIOR FILING DATE: 2000-05-04

; PRIOR APPLICATION NUMBER: 60/206,201

; PRIOR FILING DATE: 2000-05-22

; PRIOR APPLICATION NUMBER: 60/218,950

; PRIOR FILING DATE: 2000-07-14

; PRIOR APPLICATION NUMBER: 60/222,903

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: 60/223,416

; PRIOR FILING DATE: 2000-08-04

; PRIOR APPLICATION NUMBER: 60/223,378

; PRIOR FILING DATE: 2000-08-07

; NUMBER OF SEQ ID NOS: 3597

; SOFTWARE: FastSeq for Windows Version 3.0

US-10-154-884B-3110

Query Match 29.0%; Score 290; DB 12; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.2e-143;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATGTTATCCGCTCACAATCCACACA 771
Db 16 CGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATGTTATCCGCTCACAATCCACACA 75

QY 772 ACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGGTGCCTTAATGAGTGAAGTAACTCA 831
Db 76 ACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGGTGCCTTAATGAGTGAAGTAACTCA 135

QY 832 CATTAATTTGGTTGCGCTCACTGCCCGCTTCCAGTTCGCGAAACCTGTCGCGCAGCTGC 891
Db 136 CATTAATTTGGTTGCGCTCACTGCCCGCTTCCAGTTCGCGAAACCTGTCGCGCAGCTGC 195

QY 892 ATTAATGAATCGGCCAACGCGGGGAGAGCGGTTTGGTATTTGGGGCTCTTCCGCTT 951
Db 196 ATTAATGAATCGGCCAACGCGGGGAGAGCGGTTTGGTATTTGGGGCTCTTCCGCTT 255

QY 952 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
Db 256 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 305

RESULT 6

US-10-040-862-3110
; Sequence 310, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-01352005
; CURRENT APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3110
; LENGTH: 380
; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-040-862-3110

Query Match 29.0%; Score 290; DB 15; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.2e-143;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATGTTATCCGCTCACAATCCACACA 771
Db 16 CGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATGTTATCCGCTCACAATCCACACA 75

QY 772 ACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGGTGCCTTAATGAGTGAAGTAACTCA 831
Db 76 ACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGGTGCCTTAATGAGTGAAGTAACTCA 135

QY 832 CATTAATTTGGTTGCGCTCACTGCCCGCTTCCAGTTCGCGAAACCTGTCGCGCAGCTGC 891
Db 136 CATTAATTTGGTTGCGCTCACTGCCCGCTTCCAGTTCGCGAAACCTGTCGCGCAGCTGC 195

QY 892 ATTAATGAATCGGCCAACGCGGGGAGAGCGGTTTGGTATTTGGGGCTCTTCCGCTT 951
Db 196 ATTAATGAATCGGCCAACGCGGGGAGAGCGGTTTGGTATTTGGGGCTCTTCCGCTT 255

QY 952 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
Db 256 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 305

RESULT 7

US-10-270-839-103
; Sequence 103, Application US/10270839
; Publication No. US20030143586A1
; GENERAL INFORMATION:
; APPLICANT: Chao, Qimin
; APPLICANT: Grasso, Luigi
; APPLICANT: Sasse, Philip M.
; APPLICANT: Nicolaides, Nicholas C.
; TITLE OF INVENTION: Genetic Hypermutability of Plants for Gene Discovery and Diagnostics
; FILE REFERENCE: AG000205 (MOR-0133)
; CURRENT APPLICATION NUMBER: US/10/270,839
; CURRENT FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/328,750
; PRIOR FILING DATE: 2001-10-12
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 103
; LENGTH: 413
; TYPE: DNA
; ORGANISM: Conyza sp.
US-10-270-839-103

Query Match 29.0%; Score 290; DB 13; Length 413;
Best Local Similarity 100.0%; Pred. No. 1.2e-143;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATGTTATCCGCTCACAATCCACACA 771
Db 83 CGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATGTTATCCGCTCACAATCCACACA 142

QY 772 ACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGGTGCCTTAATGAGTGAAGTAACTCA 831
Db 143 ACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGGTGCCTTAATGAGTGAAGTAACTCA 202

QY 832 CATTAATTTGGTTGCGCTCACTGCCCGCTTCCAGTTCGCGAAACCTGTCGCGCAGCTGC 891
Db 203 CATTAATTTGGTTGCGCTCACTGCCCGCTTCCAGTTCGCGAAACCTGTCGCGCAGCTGC 262

QY 892 ATTAATGAATCGGCCAACGCGGGGAGAGCGGTTTGGTATTTGGGGCTCTTCCGCTT 951
Db 263 ATTAATGAATCGGCCAACGCGGGGAGAGCGGTTTGGTATTTGGGGCTCTTCCGCTT 322

QY 952 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001

QY 832 CATTAAATTCGCTGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGC 891
Db 129 CATTAAATTCGCTGCTGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGC 188
QY 892 ATTAATGAATCGGCAACGCGGGGAGAGCGGCTTTGCGTATTGGGGCTCTTCGCGTT 951
Db 189 ATTAATGAATCGGCAACGCGGGGAGAGCGGCTTTGCGTATTGGGGCTCTTCGCGTT 248
QY 952 CCTCGCTCACTGACTCGCTGCGCTGCGTTCGCTGCGCTGCGCGAGCGGTA 1001
Db 249 CCTCGCTCACTGACTCGCTGCGCTGCGTTCGCTGCGCTGCGCGAGCGGTA 298

RESULT 12

US-09-736-457-639
; Sequence 639, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 639
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-639

Query Match 29.0%; Score 290; DB 10; Length 584;
Best Local Similarity 100.0%; Pred. No. 1.2e-143;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGCTCATAGCTGTTTCTCTGCTGTAATTTGTTATCCGCTCAAAATTCACACA 771
Db 37 CGTAATCATGCTCATAGCTGTTTCTCTGCTGTAATTTGTTATCCGCTCAAAATTCACACA 96
QY 772 ACATACGAGCCGGAAGCATAAAGTGTAAAGCTGGGCTGCTTAATGAGTGAGCTAACTCA 831
Db 97 ACATACGAGCCGGAAGCATAAAGTGTAAAGCTGGGCTGCTTAATGAGTGAGCTAACTCA 156
QY 832 CATTAAATTCGCTGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGC 891
Db 157 CATTAAATTCGCTGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGC 216
QY 892 ATTAATGAATCGGCAACGCGGGGAGAGCGGCTTTGCGTATTGGGGCTCTTCGCGTT 951
Db 217 ATTAATGAATCGGCAACGCGGGGAGAGCGGCTTTGCGTATTGGGGCTCTTCGCGTT 276
QY 952 CCTCGCTCACTGACTCGCTGCGCTGCGTTCGCTGCGCTGCGCGAGCGGTA 1001
Db 277 CCTCGCTCACTGACTCGCTGCGCTGCGTTCGCTGCGCTGCGCGAGCGGTA 326

RESULT 13

US-09-902-941-639
; Sequence 639, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong

; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 639
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-941-639

Query Match 29.0%; Score 290; DB 10; Length 584;
Best Local Similarity 100.0%; Pred. No. 1.2e-143;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGCTCATAGCTGTTTCTCTGCTGTAATTTGTTATCCGCTCAAAATTCACACA 771
Db 37 CGTAATCATGCTCATAGCTGTTTCTCTGCTGTAATTTGTTATCCGCTCAAAATTCACACA 96
QY 772 ACATACGAGCCGGAAGCATAAAGTGTAAAGCTGGGCTGCTTAATGAGTGAGCTAACTCA 831
Db 97 ACATACGAGCCGGAAGCATAAAGTGTAAAGCTGGGCTGCTTAATGAGTGAGCTAACTCA 156
QY 832 CATTAAATTCGCTGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGC 891
Db 157 CATTAAATTCGCTGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGC 216
QY 892 ATTAATGAATCGGCAACGCGGGGAGAGCGGCTTTGCGTATTGGGGCTCTTCGCGTT 951
Db 217 ATTAATGAATCGGCAACGCGGGGAGAGCGGCTTTGCGTATTGGGGCTCTTCGCGTT 276
QY 952 CCTCGCTCACTGACTCGCTGCGCTGCGTTCGCTGCGCTGCGCGAGCGGTA 1001
Db 277 CCTCGCTCACTGACTCGCTGCGCTGCGTTCGCTGCGCTGCGCGAGCGGTA 326

RESULT 14

US-09-849-626-639
; Sequence 639, Application US/09849626
; Publication No. US20020197669A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya
; APPLICANT: Fanger, Gary
; APPLICANT: Wang, Aijun
; APPLICANT: Wang, Tongtong
; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 639
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-849-626-639

Query Match 29.0%; Score 290; DB 10; Length 584;

Best Local Similarity 100.0%; Pred. No. 1.2e-143;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	712	CGTAATCATGGTCATAGCTGTTCCCTGCTGTAAGAAATTTATCCGCTCAAAATTCACACA	771
Db	37	CGTAATCATGGTCATAGCTGTTCCCTGCTGTAAGAAATTTATCCGCTCAAAATTCACACA	96
QY	772	ACATAGAGCCGGAAGCATAAAGTGTAAAGCTGGGGTGCTTAATGAGTGAGCTAACTCA	831
Db	97	ACATAGAGCCGGAAGCATAAAGTGTAAAGCTGGGGTGCTTAATGAGTGAGCTAACTCA	156
QY	832	CATTAAATGCTGCTCACTGCCGCTTCCAGTCGGGAAACCTGTCGTGCCAGCTGC	891
Db	157	CATTAAATGCTGCTCACTGCCGCTTCCAGTCGGGAAACCTGTCGTGCCAGCTGC	216
QY	892	ATTAAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATTGGGCGCTCTTCCGCTT	951
Db	217	ATTAAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATTGGGCGCTCTTCCGCTT	276
QY	952	CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTA	1001
Db	277	CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTA	326

RESULT 15

US-09-476-300-639
; Sequence 639, Application US/09476300
; Publication No. US20030125245A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Banguz, Chaitanya S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C3
; CURRENT APPLICATION NUMBER: US/09/476,300
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 785
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 639
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-476-300-639

Query Match 29.0%; Score 290; DB 11; Length 584;
Best Local Similarity 100.0%; Pred. No. 1.2e-143;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	712	CGTAATCATGGTCATAGCTGTTCCCTGCTGTAAGAAATTTATCCGCTCAAAATTCACACA	771
Db	37	CGTAATCATGGTCATAGCTGTTCCCTGCTGTAAGAAATTTATCCGCTCAAAATTCACACA	96
QY	772	ACATAGAGCCGGAAGCATAAAGTGTAAAGCTGGGGTGCTTAATGAGTGAGCTAACTCA	831
Db	97	ACATAGAGCCGGAAGCATAAAGTGTAAAGCTGGGGTGCTTAATGAGTGAGCTAACTCA	156
QY	832	CATTAAATGCTGCTCACTGCCGCTTCCAGTCGGGAAACCTGTCGTGCCAGCTGC	891
Db	157	CATTAAATGCTGCTCACTGCCGCTTCCAGTCGGGAAACCTGTCGTGCCAGCTGC	216
QY	892	ATTAAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATTGGGCGCTCTTCCGCTT	951
Db	217	ATTAAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATTGGGCGCTCTTCCGCTT	276
QY	952	CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTA	1001
Db	277	CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTA	326

Search completed: February 15, 2004, 22:26:48
Job time : 423.484 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 21:50:21 ; Search time 2636.53 Seconds
(without alignments)
9227.587 Million cell updates/sec

Title: US-09-921-143-36_COPY_2000_3000
Perfect score: 1001
Sequence: 1 ttcaggctaaccagtgagg.....ttcggtcgaggagcgga 1001

Scoring table: OLIGO NUC
Gapop 60.0, Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size : 0

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST.*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estc:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est6:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	290	29.0	362	12	BI940910
2	290	29.0	391	12	BI675829
3	290	29.0	391	12	BI937546
4	290	29.0	398	9	AU224539

5	290	29.0	401	12	BI675832
6	290	29.0	418	10	BI361564
7	290	29.0	418	12	BI941233
8	290	29.0	420	28	AZ254218
9	290	29.0	424	12	BI937686
10	290	29.0	424	12	BI938078
11	290	29.0	424	12	BI938097
12	290	29.0	424	12	BI938182
13	290	29.0	424	12	BI938185
14	290	29.0	424	12	BI938325
15	290	29.0	424	12	BI938341
16	290	29.0	424	12	BI938341
17	290	29.0	427	12	BI675820
18	290	29.0	427	12	BI937679
19	290	29.0	427	12	BI938087
20	290	29.0	428	12	BI937859
21	290	29.0	429	12	BI675823
22	290	29.0	431	12	BI937767
23	290	29.0	436	28	AQ015737
24	290	29.0	438	12	BI937553
25	290	29.0	439	12	BI938342
26	290	29.0	440	12	BI941001
27	290	29.0	442	12	BI941064
28	290	29.0	446	12	BI937775
29	290	29.0	448	12	BI938167
30	290	29.0	448	12	BI938314
31	290	29.0	449	12	BI938324
32	290	29.0	450	12	BI937961
33	290	29.0	450	12	BI938090
34	290	29.0	451	14	CB865735
35	290	29.0	456	9	AU224531
36	290	29.0	456	12	BI937790
37	290	29.0	458	12	BI864963
38	290	29.0	459	14	CB876419
39	290	29.0	460	9	AW790424
40	290	29.0	461	28	B83369
41	290	29.0	462	12	BI937801
42	290	29.0	462	12	BI937979
43	290	29.0	463	14	CB859029
44	290	29.0	467	28	B82861
45	290	29.0	468	29	TA176D09P

ALIGNMENTS

RESULT 1
BI940910
LOCUS

DEFINITION

BI940910 362 bp mRNA linear EST 18-OCT-2001
df08b08.x1 Wellcome CRC PRN3 St13 17 egg animal cap Xenopus laevis
cDNA clone IMAGE:3556574.3, similar to SW:CAT4_CAMJE Q59296
CATALASE ; mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BI940910
BI940910.1 GI:16255382
EST.

Xenopus laevis (African clawed frog)
Xenopus laevis

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus.

REFERENCE

AUTHORS

Clifton, S., Johnson, S.B., Blumberg, B., Song, J., Hillier, L., Pape, D.,
Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person,
B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R.,
Waterston, R. and Wilson, R.

TITLE

Washu Xenopus EST project, 1999

JOURNAL

Unpublished

COMMENT

Contact: Sandy Clifton, Ph.D.

Washu Xenopus EST project, 1999

Washington University School of Medicine

444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu
 Library constructed by N. Garrett, E. ellefroid, and A.M. Zorn
 (Wellcome/CRC Institute). DNA Sequencing by: Washington University
 Genome Sequencing Center
 Clone distribution: Xenopus clones from this library are available
 through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
 Possible reversed clone: similarity on wrong strand
 Seq primer: -40UP from Gibco
 High quality sequence stop: 350.

FEATURES

source

1. .362
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="IMAGE:3556574"
 /tissue_type="egg, subtracted by stage 13-17 animal cap"
 /lab_host="PH10B (phage-resistant)"
 /clone_lib="Wellcome CRC PRN3 St13 17 egg animal cap"
 /note="Vector: pBSRN3; Site 1: NotI; Site 2: EcoRI; cDNAs
 were oligo-dT primed and directionally cloned. Staging
 according to Nieuwkoop and Faber. Library is subtracted
 and was constructed by N. Garrett, E. Bellefroid, and A.M.
 Zorn, (Wellcome/CRC Institute)."
 77 a 93 c 97 g 95 t

BASE COUNT

ORIGIN

Query Match 29.0%; Score 290; DB 12; Length 362;
 Best Local Similarity 100.0%; Pred. No. 4.3e-136;
 Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGGTCATAGCTGTTCTTCCTGTGTGAATTTATCCGCTCAACATTCACACA 771
 Db 69 CGTAATCATGGTCATAGCTGTTCTTCCTGTGTGAATTTATCCGCTCAACATTCACACA 128
 QY 772 ACATAGAGCGGAGCATTAAGTGTAAAGCCTGGGTGCTTAATGAGTGAAGTCACTCA 831
 Db 129 ACATAGAGCGGAGCATTAAGTGTAAAGCCTGGGTGCTTAATGAGTGAAGTCACTCA 188
 QY 832 CATTAAATGCTGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTCGCCAGCTGC 891
 Db 189 CATTAAATGCTGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTCGCCAGCTGC 248
 QY 892 ATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTTCGTAATGGCGCTTTCCGCTT 951
 Db 249 ATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTTCGTAATGGCGCTTTCCGCTT 308
 QY 952 CCTCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1001
 Db 309 CCTCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 358

RESULT 2

BI675829

LOCUS

DEFINITION BI675829 391 bp mRNA linear EST 17-SEP-2001
 dag61e05.y1 Wellcome CRC PRN3 St19 26 egg animal cap Xenopus laevis
 cDNA clone IMAGE:4785896 5' similar to TR:Q44068 Q44068
 ALPHA-HEMOLYSIN. ; mRNA sequence.

ACCESSION

BI675829

VERSION

BI675829.1

KEYWORDS

SOURCE

ORGANISM

Xenopus laevis

(African clawed frog)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

Xenopodinae; Xenopus.

1 (bases 1 to 391)

Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hillier, L., Pape, D.

, Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person

, B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R.,

Waterson, R., and Wilson, R.

WashU Xenopus EST project, 1999

Unpublished

Contact: Sandy Clifton, Ph.D.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

WashU Xenopus EST project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Library constructed by N. Garrett, E. ellefroid, and A.M. Zorn
 (Wellcome/CRC Institute). DNA Sequencing by: Washington University
 Genome Sequencing Center
 Clone distribution: Xenopus clones from this library are available
 through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
 Seq primer: -40RP from Gibco
 High quality sequence stop: 273.

FEATURES

source

1. .391
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="IMAGE:4785896"
 /tissue_type="egg, subtracted by stage 19-26 animal cap"
 /lab_host="PH10B (phage-resistant)"
 /clone_lib="Wellcome CRC PRN3 St19 26 egg animal cap"
 /note="Vector: pBSRN3; Site 1: NotI; Site 2: EcoRI; cDNAs
 were oligo-dT primed and directionally cloned. Staging
 according to Nieuwkoop and Faber. Library is subtracted
 and was constructed by N. Garrett and A.M. Zorn,
 (Wellcome/CRC Institute)."
 87 a 98 c 106 g 100 t

BASE COUNT

ORIGIN

Query Match 29.0%; Score 290; DB 12; Length 391;
 Best Local Similarity 100.0%; Pred. No. 4.4e-136;
 Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CGTAATCATGGTCATAGCTGTTCTTCCTGTGTGAATTTATCCGCTCAACATTCACACA 771
 Db 36 CGTAATCATGGTCATAGCTGTTCTTCCTGTGTGAATTTATCCGCTCAACATTCACACA 95
 QY 772 ACATAGAGCGGAGCATTAAGTGTAAAGCCTGGGTGCTTAATGAGTGAAGTCACTCA 831
 Db 96 ACATAGAGCGGAGCATTAAGTGTAAAGCCTGGGTGCTTAATGAGTGAAGTCACTCA 155
 QY 832 CATTAAATGCTGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTCGCCAGCTGC 891
 Db 156 CATTAAATGCTGCTCACTGCCGCTTTCCAGTCGGGAAACCTGTCGCCAGCTGC 215
 QY 892 ATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTTCGTAATGGCGCTTTCCGCTT 951
 Db 216 ATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTTCGTAATGGCGCTTTCCGCTT 275
 QY 952 CCTCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1001
 Db 276 CCTCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 325

RESULT 3

BI675829

LOCUS

DEFINITION

BI675829 391 bp mRNA linear EST 18-OCT-2001

dag61e05.y1 Wellcome CRC pCDNAI egg Xenopus laevis cDNA clone

IMAGE:3431758 5' similar to TR:Q44068 Q44068 ALPHA-HEMOLYSIN. ;

mRNA sequence.

ACCESSION

BI675829

VERSION

BI675829.1

KEYWORDS

SOURCE

ORGANISM

Xenopus laevis

(African clawed frog)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

Xenopodinae; Xenopus.

1 (bases 1 to 391)

Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hillier, L., Pape, D.

, Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person

, B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R.,

Waterson, R., and Wilson, R.

WashU Xenopus EST project, 1999

Unpublished

Contact: Sandy Clifton, Ph.D.

Waterston, R. and Wilson, R.
WashU Xenopus EST project, 1999
Unpublished
Contact: Sandy Clifton, Ph.D.
WashU Xenopus EST project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Library constructed by N. Garrett, P. LeMaire, A.M. Zorn, and J.B. Gurdon, (Wellcome/CRC Institute), DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: Xenopus clones from this library are available through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
Seq primer: -40RP from Gibco
High quality sequence stop: 367.

FEATURES
SOURCE

1.391
organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone IMAGE:3431758"
/tissue_type="egg"
/lab_host="PH10B (phage-res)
/clone lib="Wellcome CRC p
/note="Vector: pcDNA1; Site
were oligo-47 primed and d
constructed by N. Garrett,
Gordon (Wellcome/CRC Inst
102 g 105 g 89 g

	BASE COUNT	ORIGIN
1	0.000	0.000
2	0.000	0.000
3	0.000	0.000
4	0.000	0.000
5	0.000	0.000
6	0.000	0.000
7	0.000	0.000
8	0.000	0.000
9	0.000	0.000
10	0.000	0.000
11	0.000	0.000
12	0.000	0.000
13	0.000	0.000
14	0.000	0.000
15	0.000	0.000
16	0.000	0.000
17	0.000	0.000
18	0.000	0.000
19	0.000	0.000
20	0.000	0.000
21	0.000	0.000
22	0.000	0.000
23	0.000	0.000
24	0.000	0.000
25	0.000	0.000
26	0.000	0.000
27	0.000	0.000
28	0.000	0.000
29	0.000	0.000
30	0.000	0.000
31	0.000	0.000
32	0.000	0.000
33	0.000	0.000
34	0.000	0.000
35	0.000	0.000
36	0.000	0.000
37	0.000	0.000
38	0.000	0.000
39	0.000	0.000
40	0.000	0.000
41	0.000	0.000
42	0.000	0.000
43	0.000	0.000
44	0.000	0.000
45	0.000	0.000
46	0.000	0.000
47	0.000	0.000
48	0.000	0.000
49	0.000	0.000
50	0.000	0.000
51	0.000	0.000
52	0.000	0.000
53	0.000	0.000
54	0.000	0.000
55	0.000	0.000
56	0.000	0.000
57	0.000	0.000
58	0.000	0.000
59	0.000	0.000
60	0.000	0.000
61	0.000	0.000
62	0.000	0.000
63	0.000	0.000
64	0.000	0.000
65	0.000	0.000
66	0.000	0.000
67	0.000	0.000
68	0.000	0.000
69	0.000	0.000
70	0.000	0.000
71	0.000	0.000
72	0.000	0.000
73	0.000	0.000
74	0.000	0.000
75	0.000	0.000
76	0.000	0.000
77	0.000	0.000
78	0.000	0.000
79	0.000	0.000
80	0.000	0.000
81	0.000	0.000
82	0.000	0.000
83	0.000	0.000
84	0.000	0.000
85	0.000	0.000
86	0.000	0.000
87	0.000	0.000
88	0.000	0.000
89	0.000	0.000
90	0.000	0.000
91	0.000	0.000
92	0.000	0.000
93	0.000	0.000
94	0.000	0.000
95	0.000	0.000
96	0.000	0.000
97	0.000	0.000
98	0.000	0.000
99	0.000	0.000
100	0.000	0.000

Query Match	29.0%; Score 290; DB 12; Length 391;
Best Local Similarity	100.0%; Pred. No. 4.4e-136;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	712 CGTAATCATGCTCATAGCTGTTTCCTGTGTGAATTTGTTATCCGCTCACAATTCACACA 771
Db	12 CGTAATCATGCTCATAGCTGTTTCCTGTGTGAATTTGTTATCCGCTCACAATTCACACA 71
QY	772 ACATACGAGCCGGAGCATAAAGTGTAAACCTTGGGGTGCTTAATGATGAGCTAACTCA 831
Db	72 ACATACGAGCCGGAGCATAAAGTGTAAAGCCTTGGGGTGCTTAATGATGAGCTAACTCA 131
QY	832 CATTAATTTGCGTTTGGGCTCACTGCGCCGCTTTCCAGTCCGGGAAACCTGTGTCGCCAGCTGC 891
Db	132 CATTAATTTGCGTTTGGGCTCACTGCGCCGCTTTCCAGTCCGGGAAACCTGTGTCGCCAGCTGC 191
QY	892 ATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGGTATTTGGGCGCTCTTCCGCTT 951
Db	132 ATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGGTATTTGGGCGCTCTTCCGCTT 251
QY	952 CTTGCTCACTGACTCGCTCGGCTCGGTCGCTTCGCGTGGGGGAGCGGTA 1001
Db	252 CTTGCTCACTGACTCGCTCGGCTCGGTCGCTTCGCTCGGCGAGCGGTA 01

RESULT 4
AU224539

LOCUS	AU224539	398 bp	mRNA	linear	EST 21-OCT-2002
DEFINITION	AU224539	<i>Ipomoea trifida</i>	stigma	<i>Ipomoea trifida</i>	cDNA clone ISM-0348
ACCESSION	AU224539				
VERSION	AU224539				
KEYWORDS	EST.				
SOURCE	AU224539.1	GI:24207512			
ORGANISM	<i>Ipomoea trifida</i>				
	<i>Ipomoea trifida</i>				
	Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Convolvulaceae; Ipomoea.				
REFERENCE	1 (bases 1 to 398)				
AUTHORS	Tsuchiya,T., Ando,A., Ogawa,C., Futagami,K., Watase,T. and Koyama				

TITLE	JOURNAL	COMMENT

Expressed Sequence Tags from Reproductive Organs of *Ipomoea trifida*
cDNA Libraries
Unpublished
Contact: Tohru Tsuchiya
Faculty of Bioresources
Mie University
1515 Kamihama, Tsu, Mie 514-8507, Japan
Tel: 81-59-231-9515
Fax: 81-59-231-9515
Email: tsuchiya@bio.mie-u.ac.jp.

FEATURES
SOURCE

```

"organism"="Ipomoea trifida"
"mol_type"="rRNA"
"db_xref"="taxon:35884"
"clone"="ISM-0348"
"tissue_type"="stigma"
"dev_stage"="tr-nucleate pollen stage"
"clone_lib"="Ipomoea trifida stigma"
80 a 107 c 109 g 102 t

```

[illegible]

Query Match	29.0%;	Score 290;	DB 9;	Length 398;
Best Local Similarity	100.0%;	Pred. No. 4.4e-136;		
Matches 290;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

712	CGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCAAAATTCACACA	771
	Db	
67	CGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCAAAATTCACACA	126
	Qy	
772	ACATACGAGCCGAAGCATAAAGCTGAAAGCTGGGGTGCCTAATGAGTGAGCTAACTCA	831
	Db	
127	ACATACGAGCCGAAGCATAAAGCTGAAAGCTGGGGTGCCTAATGAGTGAGCTAACTCA	186
	Qy	
832	CATTAAATGGCTTGCCTCACTCCCGCTTTCAGTCGGGAAACCTGTCGTGCCAGCTGC	891
	Db	
187	CATTAAATGGCTTGCCTCACTCCCGCTTTCAGTCGGGAAACCTGTCGTGCCAGCTGC	246
	Qy	
892	ATTAAATGAATCGGCCAACCGCGGGGAGAGCGGTTTCGCTATTGGGGCGCTCTTCCGCTT	951
	Db	
247	ATTAAATGAATCGGCCAACCGCGGGGAGAGCGGTTTCGCTATTGGGGCGCTCTTCCGCTT	306
	Qy	
952	CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCGGACGGTA	1001
	Db	
307	CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCGGACGGTA	356

RESULT 5
BT675832

EST 17-SEP-2001
linear mRNA 401 bp
BI675832
da61g01.y1 Wellcome CRC PRN3 St19 26 egg animal cap Xenopus laevis
CDNA clone IMAGE:4785984 5' similar to TR:Q44068 Q44068
ALPHA-HEMOLYSIN, *l.*, mRNA sequence.

ACCESSION
VERSION

KEYWORDS
Xenopus laevis (African EST.
Xenopus laevis

SOURCE
ORGANISM

Xenopus laevis
Chordata; Metazoa; Eukaryota; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus.

REFERENCE
AUTHORS

Authors	Title
Carroll, S., Johnson, S.L., Blumberg, B., Song, J., Miller, J., Repetti, R., Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person, B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R., Waterston, R. and Wilson, R.	Washu Xenopus EST project, 1999

TITLE
JOURNAL:

COMMENT: Contact: Sandy Clifton, Ph.D.
WashU Xenopus EST project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St.
Louis, MO 63110

Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
Library constructed by N. Garrett, E. ellefroid, and A.M. Zorn
(Wellcome/CRC Institute). DNA Sequencing by: Washington University
Genome Sequencing Center
Clone distribution: Xenopus clones from this library are available
through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
Seq primer: -40RP from Gibco
High quality sequence stop: 291.

FEATURES

source

1. 401
Location/Qualifiers
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:4785984"
/tissue_type="egg, subtracted by stage 19-26 animal cap"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Wellcome CRC PRN3 St19 26 egg animal cap"
/note="Vector: pBSRN3; Site 1: NotI; Site 2: EcoRI; cDNAs
were oligo-dT primed and directionally cloned. Staging
according to Nieuwkoop and Faber. Library is subtracted
and was constructed by N. Garrett and A.M. Zorn,
(Wellcome/CRC Institute)."
88 a 102 c 109 g 102 t

BASE COUNT
ORIGIN

Query Match 29.0%; Score 290; DB 12; Length 401;
Best Local Similarity 100.0%; Pred. No. 4.4e-136;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 712 CGTAATCATGGTCATAGCTGTTTCCCTGTGTGAAATGTTATCCGCTCAAAATTCACACA 771
Db 46 CGTAATCATGGTCATAGCTGTTTCCCTGTGTGAAATGTTATCCGCTCAAAATTCACACA 105
QY 772 ACATAGAGCGGGAACATTAAGTGTAAAGCTGGGGTCCCTAATGAGTGAGCTAACTCA 831
Db 106 ACATAGAGCGGGAACATTAAGTGTAAAGCTGGGGTCCCTAATGAGTGAGCTAACTCA 165
QY 832 CATTAATTCGGTTCGCTCACTGCCGCTTTCAGTCGGGAAACCTGTCGCCAGCTGC 891
Db 166 CATTAATTCGGTTCGCTCACTGCCGCTTTCAGTCGGGAAACCTGTCGCCAGCTGC 225
QY 892 ATTAATGAATCGCCAAACGCGGGGAGAGCGGTTTTCGTAATTCGGGCTCTTCGGCTT 951
Db 226 ATTAATGAATCGCCAAACGCGGGGAGAGCGGTTTTCGTAATTCGGGCTCTTCGGCTT 285
QY 952 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
Db 286 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 335

RESULT 6

BF361564

LOCUS

DEFINITION
BF361564 418 bp mRNA linear EST 24-NOV-2000
pot-1-64 Differentially expressed cDNA library of optic tectum of
pigeon with removal of left retina Columba livia cDNA, mRNA
sequence.

ACCESSION

BF361564

VERSION

BF361564.1

KEYWORDS

SOURCE

ORGANISM

Columba livia (domestic pigeon)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Columbiformes; Columbidae; Columba.

1 (bases 1 to 418)

REFERENCE

AUTHORS

TITLE

A profile of differentially expressed genes in p10 pigeon optic

tacta with removal of left retina after posthatching first day by

SSH

JOURNAL

COMMENT

Unpublished

Contact: Lixia Lu

Shanghai Tiedao University Medical School
Gonghexin Rd.1238, Shanghai, 200070, China
Email: jinxia@public.sta.net.cn.

FEATURES

source

1. 418
Location/Qualifiers
/organism="Columba livia"
/mol_type="mRNA"
/db_xref="taxon:8932"
/sex="female and male"
/tissue_type="optic tecta"
/dev stage="10 day postnatal"
/clone_lib="Differentially expressed cDNA library of optic
tectum of pigeon with removal of left retina"
/note="Differentially expressed cDNA library of optic
tectum of pigeon with removal of left retina"
79 a 120 c 118 g 101 t

BASE COUNT
ORIGIN

Query Match 29.0%; Score 290; DB 10; Length 418;
Best Local Similarity 100.0%; Pred. No. 4.5e-136;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 712 CGTAATCATGGTCATAGCTGTTTCCCTGTGTGAAATGTTATCCGCTCAAAATTCACACA 771
Db 96 CGTAATCATGGTCATAGCTGTTTCCCTGTGTGAAATGTTATCCGCTCAAAATTCACACA 155
QY 772 ACATAGAGCGGGAACATTAAGTGTAAAGCTGGGGTCCCTAATGAGTGAGCTAACTCA 831
Db 156 ACATAGAGCGGGAACATTAAGTGTAAAGCTGGGGTCCCTAATGAGTGAGCTAACTCA 215
QY 832 CATTAATTCGGTTCGCTCACTGCCGCTTTCAGTCGGGAAACCTGTCGCCAGCTGC 891
Db 216 CATTAATTCGGTTCGCTCACTGCCGCTTTCAGTCGGGAAACCTGTCGCCAGCTGC 275
QY 892 ATTAATGAATCGCCAAACGCGGGGAGAGCGGTTTTCGTAATTCGGGCTCTTCGGCTT 951
Db 276 ATTAATGAATCGCCAAACGCGGGGAGAGCGGTTTTCGTAATTCGGGCTCTTCGGCTT 335
QY 952 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
Db 336 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 385

RESULT 7

BF941233

LOCUS

DEFINITION
BF941233 418 bp mRNA linear EST 18-OCT-2001
df26g12.x1 Wellcome CRC PRN3 head Xenopus laevis cDNA clone
IMAGE:3579838 3' similar to TR:Q44068 Q44068 ALPHA-HEMOLYSIN. ;,
mRNA sequence.

ACCESSION

BF941233

VERSION

BF941233.1

KEYWORDS

SOURCE

ORGANISM

Xenopus laevis (African clawed frog)

Xenopus laevis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

Xenopodinae; Xenopus.

1 (bases 1 to 418)

REFERENCE

AUTHORS

TITLE

WashU Xenopus EST project, 1999

JOURNAL

COMMENT

Contact: Sandy Clifton, Ph.D.

WashU Xenopus EST project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Library constructed by N. Garrett, K. Ryan and A.M. Zorn,

(Wellcome/CRC Institute). DNA Sequencing by: Washington University

Genome Sequencing Center
 Clone distribution: Xenopus clones from this library are available
 through the I.M.A.G.E. Consortium/LINL at: info@image.lnlnl.gov
 Possible reversed clone: similarity on wrong strand
 Seq primer: -40up from Gibco
 High quality sequence stop: 336.

FEATURES

Location/Qualifiers
 1..418
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="IMAGE:3579838"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="Wellcome CRC PRN3 head"
 /notes="Vector: pBSRN3; Site 1: NotI; Site 2: EcoRI; cDNAs
 were oligo-dT primed and directionally cloned. Staging
 according to Nieuwkoop and Faber. Library was constructed
 by N. Garrett, K. Ryan and A.M. Zorn, (Wellcome/CRC
 Institute)." 96 a 105 c 113 g 104 t

BASE COUNT
 ORIGIN

Query Match 29.0%; Score 290; DB 12; Length 418;
 Best Local Similarity 100.0%; Pred. No. 4.5e-136;
 Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 712 CGTAATCATGTCATAGCTGTTCCCTGTGTGAATTTGTTATCGCTCACAAATTCACACA 771
 Db 69 CGTAATCATGTCATAGCTGTTCCCTGTGTGAATTTGTTATCGCTCACAAATTCACACA 128
 QY 772 ACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGTCCTAATGAGTGAGTAACTCA 831
 Db 129 ACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGTCCTAATGAGTGAGTAACTCA 188
 QY 832 CATTAATGTTGCGTCTACTGCCCTTTCAGTCGGGAACCTGTGCGCACTGC 891
 Db 189 CATTAATGTTGCGTCTACTGCCCTTTCAGTCGGGAACCTGTGCGCACTGC 248
 QY 892 ATTAATGAATCGGCCAACGCGGGAGAGCGGTTTTCGTATTGGCGCTCTTCGGCTT 951
 Db 249 ATTAATGAATCGGCCAACGCGGGAGAGCGGTTTTCGTATTGGCGCTCTTCGGCTT 308
 QY 952 CTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
 Db 309 CTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 358

RESULT 8
 AZ254218 420 bp DNA linear GSS 21-JUN-2000
 LOCUS
 DEFINITION
 cg0029f RFLP sequences of cowpea, Vigna unguiculata Vigna
 unguiculata genomic, genomic survey sequence.
 ACCESSION
 AZ254218
 VERSION
 AZ254218.1 GI:8602455
 KEYWORDS
 GSS.
 SOURCE
 Vigna unguiculata (cowpea)
 ORGANISM
 Vigna unguiculata
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; rosids
 ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna.
 1 (bases 1 to 420)
 Denny,R., Danesh,D., Mudge,J., Cooper,A., Larson,K., Fatokun,C. and
 Young,N.D.
 RFLP sequences of cowpea, Vigna unguiculata
 Unpublished
 Contact: Young Nevin D
 Department of Plant Pathology
 University of Minnesota
 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul MN 55108, USA
 Tel: 612 625 2225
 Fax: 612 625 9728
 Email: nevin@tc.umn.edu

TITLE
 JOURNAL
 COMMENT

Sequence of mapped RFLP marker cg0029 on linkage group vi of
 cowpea. For more information, see BeanGenes at:
 http://ars-genome.cornell.edu/cgi-bin/WebAce/webace?db=beanGenes&cl
 ass=Locus. Please see as authorities for mapping/naming: Fatokun,
 C.A., Danesh, D., Young, N.D. (1993) RFLP linkage map for cowpea,
 (Vigna unguiculata (L.) Walp.) in: Genetic Maps, 1992. S. J. O'Brien
 , ed. Cold Spring Harbor Press, Cold Spring Harbor, NY, pp.
 6.256-6.258 - and - Menancio-Hautea, D., Kumar, L., Danesh, D.,
 Young, N.D. (1993) RFLP linkage map for mungbean (Vigna radiata (L.
) Wilczek) in: Genetic Maps, 1992. S. J. O'Brien, ed. Cold Spring
 Harbor Press, Cold Spring Harbor, NY, pp. 6.259-6.260.
 Insert Length: 400 Sd Error: 0.00
 Seq primer: M13F
 Class: RFLP probe.

FEATURES

Location/Qualifiers
 1..420
 /organism="Vigna unguiculata"
 /mol_type="genomic DNA"
 /cultivar="California Blackeye #5"
 /db_xref="taxon:3917"
 /tissue_type="Hypocotyl and roots"
 /dev_stage="Sprouts"
 /clone_lib="RFLP sequences of cowpea, Vigna unguiculata"
 /note="Vector: pUC 18; Site 1: Pst I; DNA was digested
 with Pst I, size separated by sucrose gradient
 centrifugation and the fraction between 500-3000 base
 pairs ligated into the vector using standard method." 98 a 104 c 112 g 90 t 16 others

BASE COUNT
 ORIGIN

Query Match 29.0%; Score 290; DB 28; Length 420;
 Best Local Similarity 100.0%; Pred. No. 4.5e-136;
 Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 712 CGTAATCATGTCATAGCTGTTCCCTGTGTGAATTTGTTATCGCTCACAAATTCACACA 771
 Db 20 CGTAATCATGTCATAGCTGTTCCCTGTGTGAATTTGTTATCGCTCACAAATTCACACA 79
 QY 772 ACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGTCCTAATGAGTGAGTAACTCA 831
 Db 80 ACATACGAGCGGGAAGCATAAAGTGTAAAGCTGGGTCCTAATGAGTGAGTAACTCA 139
 QY 832 CATTAATGCGTTGCGTCTACTGCCCTTTCAGTCGGGAACCTGTGCGCACTGC 891
 Db 140 CATTAATGCGTTGCGTCTACTGCCCTTTCAGTCGGGAACCTGTGCGCACTGC 199
 QY 892 ATTAATGAATCGGCCAACGCGGGAGAGCGGTTTTCGTATTGGCGCTCTTCGGCTT 951
 Db 200 ATTAATGAATCGGCCAACGCGGGAGAGCGGTTTTCGTATTGGCGCTCTTCGGCTT 259
 QY 952 CTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
 Db 260 CTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 309

RESULT 9
 BI937686 424 bp mRNA linear EST 18-OCT-2001
 LOCUS
 DEFINITION
 del3f03.x1 Wellcome CRC PRN3 dorsal lip Xenopus laevis cDNA clone
 IMAGE:3438076 3' similar to TR:Q44068 Q44068 ALPHA-HEMOLYSIN. ;
 mRNA sequence.
 ACCESSION
 BI937686
 VERSION
 BI937686.1 GI:16252159
 KEYWORDS
 EST.
 SOURCE
 Xenopus laevis (African clawed frog)
 ORGANISM
 Xenopus laevis
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 Xenopodinae; Xenopus.
 1 (bases 1 to 424)
 Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L., Pape,D.
 , Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y., Person
 ,B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,

TITLE
JOURNAL
COMMENT

Waterston, R. and Wilson, R.
WashU Xenopus EST project, 1999
Unpublished
Contact: Sandy Clifton, Ph.D.
WashU Xenopus EST project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: estowatson.wustl.edu
Library constructed by A.M. Zorn (Wellcome/CRC Institute). DNA
Sequencing by: Washington University Genome Sequencing Center
Clone distribution: Xenopus clones from this library are available
through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
Possible reversed clone: similarity on wrong strand
Seq primer: -40UP from Gibco
High quality sequence stop: 397.

FEATURES
source

1. .424
Location/Qualifiers
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:3438076"
/tissue_type="dorsal lip"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Wellcome CRC PRN3 dorsal lip"
/note="Vector: pBSRN3; Site 1: NotI; Site 2: EcoRI; cDNAs
were oligo-dT primed and directionally cloned. Staging
according to Nieuwkoop and Faber. Library was constructed
by A.M. Zorn (Wellcome/CRC Institute)."
100 a 106 c 112 g 106 t

BASE COUNT
ORIGIN

Query Match 29.0%; Score 290; DB 12; Length 424;
Best Local Similarity 100.0%; Pred. No. 4.5e-136; Indels 0; Gaps 0;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CCGTATCATGGTCATAGCTGTTCTCTGTGTGAATTTGTTATCCGCTCACAAATCCACACA 771
Db 69 CGTAATCATGGTCATAGCTGTTCTCTGTGTGAATTTGTTATCCGCTCACAAATCCACACA 128
QY 772 ACATACGAGCCGGAGCATTAAGTCTAAAGCTGGGTTGCTTAATGAGTGAAGTCACTCA 831
Db 129 ACATACGAGCCGGAGCATTAAGTCTAAAGCTGGGTTGCTTAATGAGTGAAGTCACTCA 188
QY 832 CATTAATTCGGTTGGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGCGCAGCTGC 891
Db 189 CATTAATTCGGTTGGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGCGCAGCTGC 248
QY 892 ATTAATGAATCGCCCAACGCGCGGAGAGCGGTTTCGCTATTTGGGCGCTCTTCCGCTT 951
Db 249 ATTAATGAATCGCCCAACGCGCGGAGAGCGGTTTCGCTATTTGGGCGCTCTTCCGCTT 308
QY 952 CTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
Db 309 CTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 358

RESULT 10
LOCUS

Bi938078 424 bp mRNA linear EST 18-OCT-2001
de33b12.x1 Wellcome CRC PRN3 dorsal lip Xenopus laevis cDNA clone
IMAGE:3473614 3' similar to TR:Q44068 ALPBA-HEMOLYSIN. ;
mRNA sequence.

ACCESSION

Bi938078.1 GI:16252550

VERSION

KEYWORDS

SOURCE

ORGANISM

Xenopus laevis (African clawed frog)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;

Xenopodinae; Xenopus.

REFERENCE

1 (bases 1 to 424)

AUTHORS

Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hillier, L., Pape, D.,
Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person,
B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R.,
Waterston, R. and Wilson, R.
WashU Xenopus EST project, 1999
Unpublished
Contact: Sandy Clifton, Ph.D.
WashU Xenopus EST project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: estowatson.wustl.edu
Library constructed by A.M. Zorn (Wellcome/CRC Institute). DNA
Sequencing by: Washington University Genome Sequencing Center
Clone distribution: Xenopus clones from this library are available
through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
Possible reversed clone: similarity on wrong strand
Seq primer: -40UP from Gibco
High quality sequence stop: 288.

TITLE

JOURNAL

COMMENT

FEATURES
source

1. .424
Location/Qualifiers
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:3473614"
/tissue_type="dorsal lip"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Wellcome CRC PRN3 dorsal lip"
/note="Vector: pBSRN3; Site 1: NotI; Site 2: EcoRI; cDNAs
were oligo-dT primed and directionally cloned. Staging
according to Nieuwkoop and Faber. Library was constructed
by A.M. Zorn (Wellcome/CRC Institute)."
101 a 106 c 112 g 105 t

BASE COUNT
ORIGIN

Query Match 29.0%; Score 290; DB 12; Length 424;
Best Local Similarity 100.0%; Pred. No. 4.5e-136; Indels 0; Gaps 0;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 712 CCGTATCATGGTCATAGCTGTTCTCTGTGTGAATTTGTTATCCGCTCACAAATCCACACA 771
Db 69 CGTAATCATGGTCATAGCTGTTCTCTGTGTGAATTTGTTATCCGCTCACAAATCCACACA 128
QY 772 ACATACGAGCCGGAGCATTAAGTCTAAAGCTGGGTTGCTTAATGAGTGAAGTCACTCA 831
Db 129 ACATACGAGCCGGAGCATTAAGTCTAAAGCTGGGTTGCTTAATGAGTGAAGTCACTCA 188
QY 832 CATTAATTCGGTTGGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGCGCAGCTGC 891
Db 189 CATTAATTCGGTTGGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGCGCAGCTGC 248
QY 892 ATTAATGAATCGCCCAACGCGCGGAGAGCGGTTTCGCTATTTGGGCGCTCTTCCGCTT 951
Db 249 ATTAATGAATCGCCCAACGCGCGGAGAGCGGTTTCGCTATTTGGGCGCTCTTCCGCTT 308
QY 952 CTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001
Db 309 CTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 358

RESULT 11

Bi938097

LOCUS

DEFINITION

de33f05.x1 Wellcome CRC PRN3 dorsal lip Xenopus laevis cDNA clone

IMAGE:3473792 3' similar to TR:Q44068 ALPBA-HEMOLYSIN. ;

mRNA sequence.

Bi938097.1 GI:16252569

VERSION

KEYWORDS

SOURCE

ORGANISM

Xenopus laevis (African clawed frog)

Xenopus laevis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopus laevis

1. (bases 1 to 424)

Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hillier, L., Pape, D., Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person, B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R., Waterston, R. and Wilson, R.

WashU Xenopus EST project, 1999

Unpublished

Contact: Sandy Clifton, Ph.D.

WashU Xenopus EST project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Library constructed by A.M. Zorn (Wellcome/CRC Institute). DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: Xenopus clones from this library are available through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov

Possible reversed clone: similarity on wrong strand

Seq primer: -40UP from Gibco

High quality sequence stop: 379.

Location/Qualifiers

FEATURES

source

1. .424

/organism="Xenopus laevis"

/mol_type="mRNA"

/db_xref="taxon:8355"

/clone="IMAGE:3473792"

/issue_type="dorsal lip"

/lab_host="DH10B (phage-resistant)"

/clone_lib="Wellcome CRC PRN3 dorsal lip"

/notes="Vector: pBSRN3; Site 1: NotI; Site 2: EcoRI; cDNAs were oligo-dT primed and directionally cloned. Staging according to Nieuwkoop and Faber. Library was constructed by A.M. Zorn (Wellcome/CRC Institute)."

BASE COUNT

ORIGIN

Query Match 29.0%; Score 290; DB 12; Length 424;

Best Local Similarity 100.0%; Pred. No. 4.5e-136;

Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 712 CGTAATCATGGTCATAGCTTTCTCTGTGTAATTTGTTATCCGCTCACAATTCACACA 771

Db 69 CGTAATCATGGTCATAGCTTTCTCTGTGTAATTTGTTATCCGCTCACAATTCACACA 128

Qy 772 ACATACGAGCGGAGCATAAAGTGAAGCTGGGCTGCTAATGAGTGAGCTAACTCA 831

Db 129 ACATACGAGCGGAGCATAAAGTGAAGCTGGGCTGCTAATGAGTGAGCTAACTCA 188

Qy 832 CATTAATTCGGTTCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTGTCAGCTGC 891

Db 189 CATTAATTCGGTTCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTGTCAGCTGC 248

Qy 892 ATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATGGCGCTCTTCCGCTT 951

Db 249 ATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATGGCGCTCTTCCGCTT 308

Qy 952 CCTCGCTCACTGACTCGCTCGGCTCGGCTCGGCTCGGCTCGGCGAGCGGTA 1001

Db 309 CCTCGCTCACTGACTCGCTCGGCTCGGCTCGGCTCGGCTCGGCGAGCGGTA 358

RESULT 12

Bi938182

LOCUS

DEFINITION de35g11.x1 Wellcome CRC PRN3 dorsal lip Xenopus laevis cDNA clone IMAGE:3473792 3' similar to TR:Q44068 Q44068 ALPHA-HEMOLYSIN. ; mRNA sequence.

ACCESSION

Bi938182

VERSION

G1:16252654

KEYWORDS

EST.

SOURCE

ORGANISM

Xenopus laevis (African clawed frog)

Xenopus laevis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae; Xenopus.

REFERENCE

AUTHORS

1. (bases 1 to 424)

Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hillier, L., Pape, D., Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person, B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R., Waterston, R. and Wilson, R.

TITLE

JOURNAL

COMMENT

WashU Xenopus EST project, 1999

Unpublished

Contact: Sandy Clifton, Ph.D.

WashU Xenopus EST project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Library constructed by A.M. Zorn (Wellcome/CRC Institute). DNA

Sequencing by: Washington University Genome Sequencing Center

Clone distribution: Xenopus clones from this library are available

through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov

Possible reversed clone: similarity on wrong strand

Seq primer: -40UP from Gibco

High quality sequence stop: 309.

Location/Qualifiers

source

1. .424

/organism="Xenopus laevis"

/mol_type="mRNA"

/db_xref="taxon:8355"

/clone="IMAGE:3473876"

/issue_type="dorsal lip"

/lab_host="DH10B (phage-resistant)"

/clone_lib="Wellcome CRC PRN3 dorsal lip"

/note="Vector: pBSRN3; Site 1: NotI; Site 2: EcoRI; cDNAs were oligo-dT primed and directionally cloned. Staging according to Nieuwkoop and Faber. Library was constructed by A.M. Zorn (Wellcome/CRC Institute)."

BASE COUNT

ORIGIN

Query Match 29.0%; Score 290; DB 12; Length 424;

Best Local Similarity 100.0%; Pred. No. 4.5e-136;

Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 712 CGTAATCATGGTCATAGCTTTCTCTGTGTAATTTGTTATCCGCTCACAATTCACACA 771

Db 69 CGTAATCATGGTCATAGCTTTCTCTGTGTAATTTGTTATCCGCTCACAATTCACACA 128

Qy 772 ACATACGAGCGGAGCATAAAGTGAAGCTGGGCTGCTAATGAGTGAGCTAACTCA 831

Db 129 ACATACGAGCGGAGCATAAAGTGAAGCTGGGCTGCTAATGAGTGAGCTAACTCA 188

Qy 832 CATTAATTCGGTTCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTGTCAGCTGC 891

Db 189 CATTAATTCGGTTCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTGTCAGCTGC 248

Qy 892 ATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATGGCGCTCTTCCGCTT 951

Db 249 ATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTGGTATGGCGCTCTTCCGCTT 308

Qy 952 CCTCGCTCACTGACTCGCTCGGCTCGGCTCGGCTCGGCTCGGCGAGCGGTA 1001

Db 309 CCTCGCTCACTGACTCGCTCGGCTCGGCTCGGCTCGGCTCGGCGAGCGGTA 358

RESULT 13

Bi938185

LOCUS

DEFINITION de35h10.x1 Wellcome CRC PRN3 dorsal lip Xenopus laevis cDNA clone IMAGE:3473922 3' similar to TR:Q44068 Q44068 ALPHA-HEMOLYSIN. ; mRNA sequence.

Bi938185

424 bp

mRNA

linear

EST 18-OCT-2001

ACCESSION BI938185
 VERSION BI938185.1 GI:16252657
 KEYWORDS EST.
 SOURCE Xenopus laevis (African clawed frog)
 ORGANISM Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 Xenopodinae; Xenopus.
 REFERENCE 1 (bases 1 to 424)
 AUTHORS Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hillier, L., Pape, D.,
 Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person,
 B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R.,
 Waterston, R. and Wilson, R.
 TITLE Washu Xenopus EST project, 1999
 JOURNAL Unpublished
 COMMENT Contact: Sandy Clifton, Ph.D.
 Washu Xenopus EST project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Library constructed by A.M. Zorn (Wellcome/CRC Institute). DNA
 Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: Xenopus clones from this library are available
 through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
 Possible reversed clone: similarity on wrong strand
 Seq primer: -40UP from Gibco
 High quality sequence stop: 375.

FEATURES
 source
 1..424
 Location/Qualifiers
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="IMAGE:3473922"
 /tissue_type="dorsal lip"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="Wellcome CRC PRN3 dorsal lip"
 /note="vector: pBSRN3; Site 1: NotI; Site 2: EcoRI; cDNAs
 were oligo-dT primed and directionally cloned. Staging
 according to Nieuwkoop and Faber. Library was constructed
 by A.M. Zorn (Wellcome/CRC Institute)."
 BASE COUNT 100 a 106 c 113 g 105 t

Query Match 29.0%; Score 290; DB 12; Length 424;
 Best Local Similarity 100.0%; Pred. No. 4.5e-136;
 Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 712 CGTAATCATGTCATAGCTGTTTCTGTGTAATTTGTTATCCGCTCACAAATCCACACA 771
 Db 69 CGTAATCATGTCATAGCTGTTTCTGTGTAATTTGTTATCCGCTCACAAATCCACACA 128
 QY 772 ACATACGACCGGAGCATTAAGCTTAAGCTCGGGTGCTTAATGAGTGAAGTCACTCA 831
 Db 129 ACATACGACCGGAGCATTAAGCTTAAGCTCGGGTGCTTAATGAGTGAAGTCACTCA 188
 QY 832 CATTAATTCGCTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTCAGCTGC 891
 Db 189 CATTAATTCGCTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTCAGCTGC 248
 QY 892 ATTATGAATCGGCCAACCGCGGGGAGAGCGGTTTCGTTATGGGCGCTCTTCCGCTT 951
 Db 249 ATTATGAATCGGCCAACCGCGGGGAGAGCGGTTTCGTTATGGGCGCTCTTCCGCTT 308
 QY 952 CCTCGCTCACTGACTCGCTCGGCTCGGCTCGGCTCGGCTCGGCGGCGGTA 1001
 Db 309 CCTCGCTCACTGACTCGCTCGGCTCGGCTCGGCTCGGCTCGGCGGCGGTA 358

RESULT 14
 BI938325
 LOCUS BI938325 424 bp mRNA linear EST 18-OCT-2001

DEFINITION dea1b11.x1 Wellcome CRC PRN3 dorsal lip Xenopus laevis cDNA clone
 IMAGE:3474380 3' similar to TR:Q44068 Q44068 ALPHA-HEMOLYSIN. ;
 mRNA sequence.
 ACCESSION BI938325
 VERSION BI938325.1 GI:16252797
 KEYWORDS EST.
 SOURCE Xenopus laevis (African clawed frog)
 ORGANISM Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 Xenopodinae; Xenopus.
 REFERENCE 1 (bases 1 to 424)
 AUTHORS Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hillier, L., Pape, D.,
 Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person,
 B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R.,
 Waterston, R. and Wilson, R.
 TITLE Washu Xenopus EST project, 1999
 JOURNAL Unpublished
 COMMENT Contact: Sandy Clifton, Ph.D.
 Washu Xenopus EST project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Library constructed by A.M. Zorn (Wellcome/CRC Institute). DNA
 Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: Xenopus clones from this library are available
 through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
 Possible reversed clone: similarity on wrong strand
 Seq primer: -40UP from Gibco
 High quality sequence stop: 397.

FEATURES

source
 1..424
 Location/Qualifiers
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="IMAGE:3474380"
 /tissue_type="dorsal lip"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="Wellcome CRC PRN3 dorsal lip"
 /note="vector: pBSRN3; Site 1: NotI; Site 2: EcoRI; cDNAs
 were oligo-dT primed and directionally cloned. Staging
 according to Nieuwkoop and Faber. Library was constructed
 by A.M. Zorn (Wellcome/CRC Institute)."
 BASE COUNT 100 a 106 c 112 g 105 t

ORIGIN

Query Match 29.0%; Score 290; DB 12; Length 424;
 Best Local Similarity 100.0%; Pred. No. 4.5e-136;
 Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 712 CGTAATCATGTCATAGCTGTTTCTGTGTAATTTGTTATCCGCTCACAAATCCACACA 771
 Db 69 CGTAATCATGTCATAGCTGTTTCTGTGTAATTTGTTATCCGCTCACAAATCCACACA 128
 QY 772 ACATACGACCGGAGCATTAAGCTTAAGCTCGGGTGCTTAATGAGTGAAGTCACTCA 831
 Db 129 ACATACGACCGGAGCATTAAGCTTAAGCTCGGGTGCTTAATGAGTGAAGTCACTCA 188
 QY 832 CATTAATTCGCTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTCAGCTGC 891
 Db 189 CATTAATTCGCTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTCAGCTGC 248
 QY 892 ATTATGAATCGGCCAACCGCGGGGAGAGCGGTTTCGTTATGGGCGCTCTTCCGCTT 951
 Db 249 ATTATGAATCGGCCAACCGCGGGGAGAGCGGTTTCGTTATGGGCGCTCTTCCGCTT 308
 QY 952 CCTCGCTCACTGACTCGCTCGGCTCGGCTCGGCTCGGCTCGGCGGCGGTA 1001
 Db 309 CCTCGCTCACTGACTCGCTCGGCTCGGCTCGGCTCGGCTCGGCGGCGGTA 358

```

RESULT 15
BI938341
LOCUS
DEFINITION
  424 bp mRNA linear EST 18-OCT-2001
  de41f04.x1 Wellcome CRC PRN3 dorsal lip Xenopus laevis cDNA clone
  IMAGE:3474558 3' similar to TR:Q44068 Q44068 ALPHA-HEMOLYSIN. ;,
  mRNA sequence.
ACCESSION
BI938341
VERSION
BI938341.1 GI:16252813
KEYWORDS
EST.
SOURCE
Xenopus laevis (African clawed frog)
ORGANISM
Xenopus laevis
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
  Xenopodinae; Xenopus.
  1 (bases 1 to 424)
REFERENCE
  Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L., Pape,D.,
  Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y., Person
  ,B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,
  Waterston,R. and Wilson,R.
  Washu Xenopus EST project, 1999
  Unpublished
  Contact: Sandy Clifton, Ph.D.
  Washu Xenopus EST project, 1999
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: est@watson.wustl.edu
  Library constructed by A.M. Zorn (Wellcome/CRC Institute). DNA
  Sequencing by: Washington University Genome Sequencing Center
  Clone distribution: Xenopus clones from this library are available
  through the I.M.A.G.E. Consortium/LNL at: info@image.llnl.gov
  Possible reversed clone: similarity on wrong strand
  Seq primer: -40UP from Gibco
  High quality sequence stop: 397.
  Location/Qualifiers
    1..424
      /organism="Xenopus laevis"
      /mol_type="mRNA"
      /db_xref="taxon:8355"
      /clone="IMAGE:3474558"
      /tissue_type="dorsal lip"
      /lab_host="DH10B (phage-resistant)"
      /clone_lib="Wellcome CRC PRN3 dorsal lip"
      /note="Vector: pBSRN3, Site 1: NotI, Site 2: EcoRI; cDNAs
      were oligo-dT primed and directionally cloned. Staging
      according to Nieuwkoop and Faber. Library was constructed;
      by A.M. Zorn (Wellcome/CRC Institute). "
      by A.M. Zorn 113 g 106 t
BASE COUNT
  99 a 106 c 113 g 106 t
ORIGIN
  Query March 29 0%; Score 290; DB 12; Length 424;
  Best Local Similarity 100.0%; Pred. No. 4.5e-136; Indels 0; Gaps 0;
  Matches 290; Conservative 0; Mismatches 0;
  QY 712 CGTAATCATGGTCATAGCTGTTTCCTGTGTGAATTTGTTATCCGCTCACAAATCCACACA 771
  Db 69 CGTAATCATGGTCATAGCTGTTTCCTGTGTGAATTTGTTATCCGCTCACAAATCCACACA 128
  QY 772 ACATCGAGCCGGAACNTAAAGTGTAAAGCTGGGTGGCTTATGAGTGAGCTAACTCA 831
  Db 129 ACATCGAGCCGGAACNTAAAGTGTAAAGCTGGGTGGCTTATGAGTGAGCTAACTCA 189
  QY 832 CATTAAATGCGTTGCGCTCACTGCCCGCTTTCAGTCCGGAAACCTGTCGTCAGCTGC 891
  Db 189 CATTAAATGCGTTGCGCTCACTGCCCGCTTTCAGTCCGGAAACCTGTCGTCAGCTGC 248
  QY 892 ATTAATGATCGCCCAACCGCGGAGAGCGGTTTGGCTATTGGGGCTCTTCCGCTT 951
  Db 249 ATTAATGATCGCCCAACCGCGGAGAGCGGTTTGGCTATTGGGGCTCTTCCGCTT 309
  QY 952 CCTCGCTCACTGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1001

```

```

Db 309 CCTCGCTCACTGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTA 358
Search completed: February 16, 2004, 03:15:45
Job time : 2640.53 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 19:10:26 ; Search time 3961.99 Seconds
(without alignments)
10335.846 Million cell updates/sec

Title: US-09-921-143-36_COPY_4000_5000

Perfect score: 1001

Sequence: 1 cgcgtggagatcatccagcc.....gccatcagatccttggcggc 1001

Scoring table:

CLICO NUC

Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 20454813386 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.nam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.hgo.hum.*

40: em.hgo.mus.*

41: em.hgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	1001	100.0	1300	1	ISTNSX	V00618 Transposon
C 2	1001	100.0	1300	6	I01012	I01012 Sequence 4
C 3	1001	100.0	1500	6	E02455	E02455 DNA encodin
C 4	1001	100.0	2220	6	AR067785	AR067785 Sequence
C 5	1001	100.0	2320	12	SYNKANRA	M84115 Cloning vec
C 6	1001	100.0	2356	1	ECU32991	U32991 Escherichia
C 7	1001	100.0	3267	12	AY048743	AY048743 Template
C 8	1001	100.0	3740	6	A00782	A00782 Plasmid pDM
C 9	1001	100.0	3740	6	A02231	A02297 Plasmid pDM
C 10	1001	100.0	3740	6	A02297	A07054 Artificial
C 11	1001	100.0	3740	6	A07054	A10362 Artificial
C 12	1001	100.0	3740	6	A10362	A14595 Synthetic n
C 13	1001	100.0	3740	6	A14595	A19608 Artificial
C 14	1001	100.0	3740	6	A19608	AR096518 Sequence
C 15	1001	100.0	3740	6	A30452	AR098395 Sequence
C 16	1001	100.0	3974	6	AR096518	AR098395 Sequence
C 17	1001	100.0	3974	6	AR098395	AR099244 Sequence
C 18	1001	100.0	3974	6	AR099244	AR172638 Sequence
C 19	1001	100.0	3974	6	AR172638	AR207786 Sequence
C 20	1001	100.0	3974	6	AR207786	AR218546 Sequence
C 21	1001	100.0	3974	6	AR218546	AR253199 Sequence
C 22	1001	100.0	3974	6	AR253199	AR256379 Sequence
C 23	1001	100.0	3974	6	AR256379	AR267137 Sequence
C 24	1001	100.0	3974	6	AR267137	BD022735 Mammary c
C 25	1001	100.0	3974	6	BD022735	BD033415 Astinicro
C 26	1001	100.0	3974	6	BD033415	BD129548 Polynucle
C 27	1001	100.0	3974	6	BD129548	BD131214 Apoptosis
C 28	1001	100.0	3974	6	BD131214	BD134158 Vascular
C 29	1001	100.0	3974	6	BD134158	BD134899 Apoptosis
C 30	1001	100.0	3974	6	BD134899	BD136632 Therapeut
C 31	1001	100.0	3974	6	BD136632	AR134670 Sequence
C 32	1001	100.0	3984	6	AR134670	AR219194 Sequence
C 33	1001	100.0	3984	6	AR219194	AX589681 Sequence
C 34	1001	100.0	3997	6	AX589681	AR134669 Sequence
C 35	1001	100.0	4208	6	AR134669	AR219193 Sequence
C 36	1001	100.0	4208	6	AR219193	AR256415 Sequence
C 37	1001	100.0	4208	6	AR256415	BD005772 Therapeut
C 38	1001	100.0	4208	6	BD005772	AR207828 Sequence
C 39	1001	100.0	4256	6	AR207828	AR134671 Sequence
C 40	1001	100.0	4277	6	AR134671	A92866 Sequence 2
C 41	1001	100.0	4277	6	A92866	AR137192 Sequence
C 42	1001	100.0	4328	6	AR137192	AR158346 Sequence
C 43	1001	100.0	4328	6	AR158346	AR241208 Sequence
C 44	1001	100.0	4328	6	AR241208	
C 45	1001	100.0	4328	6	AR241208	

ALIGNMENTS

RESULT 1
ISTNSX/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

ISTNSX 1300 bp DNA linear BCT 02-SBP-1999
Transposon Tn5 fragment encoding neomycin and kanamycin resistance
(neo) and a fragment of the reading frame of a further protein.

V00618 J01834

GI:43748

drug resistance gene; neomycin resistance; transferase.

Escherichia coli

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

Enterobacteriaceae; Escherichia

1 (bases 1 to 1300)

Beck, E., Ludwig, G., Auerwald, E.A., Reiss, B. and Schaller, H.

Nucleotide sequence and exact localization of the neomycin

Db 904 GCGATCGGAGCGCGATACCGTAAAGCAGAGAGCGGTGAGCCCATTCGCGGCGAA 845
QY 241 GCTCTTCAGCAATATCAGGGTAGCAACGCTATGTCTGTATAGCGGTCCGCCACACCA 300
Db 844 GCTCTTCAGCAATATCAGGGTAGCAACGCTATGTCTGTATAGCGGTCCGCCACACCA 785
QY 301 GCGGCGCAGTCGATGATCCAGAAAGCGCCATTTCCACCATGATATTCGGCAGC 360
Db 784 GCGGCGCAGTCGATGATCCAGAAAGCGCCATTTCCACCATGATATTCGGCAGC 725
QY 361 AGGCATCGCATGGGTACAGAGATCTCGCGGTGCGGATGCGGCTTTGAGCCTGG 420
Db 724 AGGCATCGCATGGGTACAGAGATCTCGCGGTGCGGATGCGGCTTTGAGCCTGG 665
QY 421 CGAAAGTTCGGTCCGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 480
Db 664 CGAAAGTTCGGTCCGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 605
QY 481 GACCGGCTTCCATCCGATGATGCTGCTCGATGCGATGTTTGGTGGTGGATG 540
Db 604 GACCGGCTTCCATCCGATGATGCTGCTCGATGCGATGTTTGGTGGTGGATG 545
QY 541 GCGAGTACCGGATCAAGGTATGACGCGCGCATTTGATAGCCCATGATGATCTT 600
Db 544 GCGAGTACCGGATCAAGGTATGACGCGCGCATTTGATAGCCCATGATGATCTT 485
QY 601 TCTCGGAGGACGAGGTGAGTACAGAGATCTGCGCGGCGCATTCGCCCATAGCA 660
Db 484 TCTCGGAGGACGAGGTGAGTACAGAGATCTGCGCGGCGCATTCGCCCATAGCA 425
QY 661 GCGAGTCCCTTCCGCTTCAGTGACAGCTGACAGCATGCTGCGCAAGGACGCGCGT 720
Db 424 GCGAGTCCCTTCCGCTTCAGTGACAGCTGACAGCATGCTGCGCAAGGACGCGCGT 365
QY 721 TGGCAGGACGATAGCGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 364 TGGCAGGACGATAGCGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 305
QY 781 CGGTCTTGACAAAAGAACCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 304 CGGTCTTGACAAAAGAACCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 245
QY 841 AGCAGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 244 AGCAGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 185
QY 901 GAGAACCTCGGTGCAATCCATCTGTTCAATCATGCGGAAAGCATCTCTCTCTT 960
Db 184 GAGAACCTCGGTGCAATCCATCTGTTCAATCATGCGGAAAGCATCTCTCTCTT 125
QY 961 GATCAGATCTTGCATCCCTCGCCATCAGATCTCTTGGCGG 1001
Db 124 GATCAGATCTTGCATCCCTCGCCATCAGATCTCTTGGCGG 84

RESULT 3
E02455/c
LOCUS E02455 1600 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA encoding aminoglycoside 3'-phosphotransferaseII (APH).
ACCESSION E02455
VERSION E02455 1 GI:2170686
KEYWORDS JP 1990150282-A/1.
SOURCE Transposon Tn5
ORGANISM Transposon Tn5
REFERENCE 1 (bases 1 to 1600)
AUTHORS Yoshida,N., Kikuchi,N., Shin,M. and Teraoka,H.
TITLE MODIFIED HUMAN PSTI
JOURNAL Patent: JP 1990150282-A 1 08-JUN-1990;
SHIONOGI & CO LTD
OS Transposon Tn5
PN JP 1990150282-A/1
PD 08-JUN-1990

PF 11-OCT-1988 JP 1988255580
PR 19-JUL-1988 JP 88P 181316
PI YOSHIDA NOBUO, KIKUCHI NORIHISA, SHIN MASARU, TERAOKA HIROSHI
PC C12N15/12,C07K7/10//C07H21/04,C12P21/02,C12P21/02,C12R1:19,
C07K9:00;
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: clone=pNEO;
FH Key Location/Qualifiers
FT 5'UTR 1..450
FT CDS 451..1245
FT /product='APH'
FT mat_peptide 451..1242
FT /product='APH'
FT 3'UTR 1246..1600.
FT Location/Qualifiers
source 1..1600
/organism='Transposon Tn5'
/mol_type='genomic DNA'
/db_xref='taxon:2411'
BASE COUNT 324 a 442 c 508 g 326 t
ORIGIN
Query Match 100.0%; Score 1001; DB 6; Length 1600;
Best Local Similarity 100.0%; Freq. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGTGAGGATCATCAGCGCGGTCCCGAAAAAGATTCCGAAGCCCAACCTTTTCATAG 60
Db 1384 CGTGTGAGGATCATCAGCGCGGTCCCGAAAAAGATTCCGAAGCCCAACCTTTTCATAG 1325
QY 61 AAGCGCGGTGATTCGAATCTGTGATGCGAGGTTGGGCGTGTGCTGCGGTCAAT 120
Db 1324 AAGCGCGGTGATTCGAATCTGTGATGCGAGGTTGGGCGTGTGCTGCGGTCAAT 1265
QY 121 TCGAACCCAGAGTCCGCTCAGAAAGACTCTCTCAAGAGGCGATAGAGGCGATCGCT 180
Db 1264 TCGAACCCAGAGTCCGCTCAGAAAGACTCTCTCAAGAGGCGATAGAGGCGATCGCT 1205
QY 181 GCGAATCGGAGCGGCGATACCGTAAAGCAGAGGAGCGGTGAGCCCATTCGCGCCAA 240
Db 1204 GCGAATCGGAGCGGCGATACCGTAAAGCAGAGGAGCGGTGAGCCCATTCGCGCCAA 1145
QY 241 GCTCTTCAGCAATATCAGCGGTAGCCAGCTATGCTCTGATAGCGGTCCGCCACCA 300
Db 1144 GCTCTTCAGCAATATCAGCGGTAGCCAGCTATGCTCTGATAGCGGTCCGCCACCA 1085
QY 301 GCGGCGCAGTCGATGATCCAGAAAGCGGCCATTTTCCACCATGATATTCGCGAAGC 360
Db 1084 GCGGCGCAGTCGATGATCCAGAAAGCGGCCATTTTCCACCATGATATTCGCGAAGC 1025
QY 361 AGGCATCGCATGGGTACAGAGATCTCGCGGTGCGGATGCGGCTTTGAGCCTGG 420
Db 1024 AGGCATCGCATGGGTACAGAGATCTCGCGGTGCGGATGCGGCTTTGAGCCTGG 965
QY 421 CGAAAGTTCGGTCCGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 480
Db 964 CGAAAGTTCGGTCCGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 905
QY 481 GACCGGCTTCCATTCGAGTACCGTCTCGATGCGATGTTTGGTGGTGGATG 540
Db 904 GACCGGCTTCCATTCGAGTACCGTCTCGATGCGATGTTTGGTGGTGGATG 845
QY 541 GCGAGTACCGGATCAAGCGGTATGCGCGCGCATTCATCAGCATGATGATGATCTT 600
Db 844 GCGAGTACCGGATCAAGCGGTATGCGCGCGCATTCATCAGCATGATGATGATCTT 785
QY 601 TCTCGGAGGACGAGGTGAGTACAGAGATCTCTGCGCGGCGCATCTTCCCGCAATAGCA 660
Db 784 TCTCGGAGGACGAGGTGAGTACAGAGATCTCTGCGCGGCGCATCTTCCCGCAATAGCA 725

QY 661 GCAGTCCCTTCCGCTTCACTGACAACTCGAGCAGCTGGCGAAGAACGCCGCTCG 720
Db 724 GCAGTCCCTTCCGCTTCACTGACAACTCGAGCAGCTGGCGAAGAACGCCGCTCG 665
QY 721 TGGCCAGCCAGATAGCCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 664 TGGCCAGCCAGATAGCCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 605
QY 781 CGGTCTTGACAAAAGAACCGGCGCCCTGACGACGCGGACGCGGACGCGATCAG 840
Db 604 CGGTCTTGACAAAAGAACCGGCGCCCTGACGACGCGGACGCGGACGCGATCAG 545
QY 841 AGCAGCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 544 AGCAGCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 485
QY 901 GAGAACCTGGTGAATCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 484 GAGAACCTGGTGAATCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 425
QY 961 GATCAGATCTTGATCCCTCGGCGCATCAGATCCTTGGCGGC 1001
Db 424 GATCAGATCTTGATCCCTCGGCGCATCAGATCCTTGGCGGC 384

RESULT 4

LOCUS AR067785 2220 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5851808.
ACCESSION AR067785
VERSION AR067785.1 GI:5999007
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2220)
AUTHORS Elledge, S.J. and Liu, Q.
TITLE Rapid subcloning using site-specific recombination
JOURNAL Patent: US 5851808-A 1 22-DEC-1998;
FEATURES Location/Qualifiers
source 1..2220
/organism="unknown"
BASE COUNT 506 a 610 c 581 g 523 t
ORIGIN

Query Match 100.0%; Score 1001; DB 6; Length 2220;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGCTGGAGGATCATCCAGCGCGGTCCCGGAAACGATCCGAAAGCCCAACCTTTCATAG 60
Db 932 CGCTGGAGGATCATCCAGCGCGGTCCCGGAAACGATCCGAAAGCCCAACCTTTCATAG 991
QY 61 AAGCGCGCGTGGAAATCGAATCTCGTGATGGCAGGTTGGCGTGCCTGCTGCTGCTGCT 120
Db 992 AAGCGCGCGTGGAAATCGAATCTCGTGATGGCAGGTTGGCGTGCCTGCTGCTGCTGCT 1051
QY 121 TCGAACCCACAGTCCCGCTCAGAGAACCTCGTCAAGAGCGGATAGAGGCGATCGCT 180
Db 1052 TCGAACCCACAGTCCCGCTCAGAGAACCTCGTCAAGAGCGGATAGAGGCGATCGCT 1111
QY 181 CGCAATCGGAGCGCGGATACCGTAAAGCAGGAGCGGTGAGCGGATTCGCCGCCAA 240
Db 1112 CGCAATCGGAGCGCGGATACCGTAAAGCAGGAGCGGTGAGCGGATTCGCCGCCAA 1171
QY 241 GCTCTTCAGCATATCAGGATAGCGGATAGCGGATAGCGGATAGCGGATAGCGGATAG 300
Db 1172 GCTCTTCAGCATATCAGGATAGCGGATAGCGGATAGCGGATAGCGGATAGCGGATAG 1231
QY 301 GCCGCCACAGTCCGATGAATCCAGAAAGCGGCGATTTCCACATGATATTCGCGAAGC 360
Db 1232 GCCGCCACAGTCCGATGAATCCAGAAAGCGGCGATTTCCACATGATATTCGCGAAGC 1291

QY 361 AGCATCGCATGGGTCAAGCAGAGATCCTCGCGCTGGGCATCGCGCCTTGAGCCTGG 420
Db 1292 AGCATCGCATGGGTCAAGCAGAGATCCTCGCGCTGGGCATCGCGCCTTGAGCCTGG 1351
QY 421 CGAACAGTTCGGCTTGGCGCGAGCCCTGATGCTCTTCGTCAGATCATCTGATCGACAA 480
Db 1352 CGAACAGTTCGGCTTGGCGCGAGCCCTGATGCTCTTCGTCAGATCATCTGATCGACAA 1411
QY 481 GACCGGCTTCATCCGAGTACGCTCGCTCGATCGATGCTTTCGCTTGGTGGTCAATG 540
Db 1412 GACCGGCTTCATCCGAGTACGCTCGCTCGATCGATGCTTTCGCTTGGTGGTCAATG 1471
QY 541 GCGAGGTAGCGGATCAAGCGTATGACGCGCGCATTCGATCAGCCATCATGATGATACCT 600
Db 1472 GCGAGGTAGCGGATCAAGCGTATGACGCGCGCATTCGATCAGCCATCATGATGATACCT 1531
QY 601 TCTCGCAGAGCAAGGTGATGACAGAGATCTCTCCCGCGGCACTTCGCCCATAGCA 660
Db 1532 TCTCGCAGAGCAAGGTGATGACAGAGATCTCTCCCGCGGCACTTCGCCCATAGCA 1591
QY 661 GCCAGTCCCTTCCCGCTTCACTGACAACTCGAGCAGCTGGCAGAGAACGCCCGCTCG 720
Db 1592 GCCAGTCCCTTCCCGCTTCACTGACAACTCGAGCAGCTGGCAGAGAACGCCCGCTCG 1651
QY 721 TGGCCAGCCAGATAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 1652 TGGCCAGCCAGATAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1711
QY 781 CGGTCTTGACAAAAGAACCGGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 1712 CGGTCTTGACAAAAGAACCGGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1771
QY 841 AGCAGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 1772 AGCAGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1831
QY 901 GAGAACCTGGTGAATCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 1832 GAGAACCTGGTGAATCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1891
QY 961 GATCAGATCTTGATCCCTCGGCGCATCAGATCCTTGGCGGC 1001
Db 1892 GATCAGATCTTGATCCCTCGGCGCATCAGATCCTTGGCGGC 1932

RESULT 5

LOCUS SYNKANRA
DEFINITION Cloning vector pHS96 kanamycin resistance protein (neo) gene,
putative cds.
ACCESSION M84115
VERSION M84115.1 GI:146531
KEYWORDS kanamycin resistance,
unidentified cloning vector
SOURCE unidentified cloning vector
ORGANISM artificial sequences; vectors.
REFERENCE 1 (sites)
AUTHORS Seifert, H.S., So, M. and Heffron, F.
TITLE Shuttle mutagenesis: a method of introducing transposons into
transformable organisms
JOURNAL Genet. Eng. Prin. Methods 8, 327-336 (1986)
REFERENCE 2 (bases 1 to 2320)
AUTHORS Seifert, H.S., Chen, E.Y., So, M. and Heffron, F.
TITLE Shuttle mutagenesis: a method of transposon mutagenesis for
Saccharomyces cerevisiae
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 83 (3), 735-739 (1986)
MEDLINE 86121014
PUBMED 3003748
COMMENT Original source text: Cloning vector (sub_species Cloning vector
pHS96) DNA.
Cloning vector pHS96 is used in the system of shuttle mutagenesis.
FEATURES Location/Qualifiers

Adelaide, S.A. 5001, Australia
 On Jan 23, 1996 this sequence version replaced gi:1000122.
 COMMENT Location/Qualifiers
 FEATURES
 source
 1. .2356
 /organism="Escherichia coli"
 /mol_type="genomic DNA"
 /strain="SM10 (lambda pir)"
 /db_xref="taxon:562"
 /plasmid="mini-Tn5 Km"
 /transposon="mini-Tn5 Km"
 repeat_region
 1. .19
 /note="Tn5 I end functional in transposition"
 /citations=[4]
 /rpt_type=inverted
 857..1651
 CDS
 /functions="neomycin and kanamycin resistance"
 /codon_start=11
 /transl_table=11
 /product="neomycin phosphotransferase"
 /protein_id="AA85506.1"
 /db_xref="GI:1000123"
 /translations="MIEQDGLHAGSPAAWVERLFGYDWAQQTIGCSAAVRLSAQR
 LSHLAPEKVSIMADAMERLHTLDPATCFDFOAKHRIERARTRMEAGLVDDDLDE
 EHOGLAPAEIPARKARMEDGDLVTHGDACLEIMVENGEPSPIDGRLGVADRY
 QDALATRDIAELGSEWADRFVLVYGLIAPDSQRIAYRLLEFF"
 repeat_region
 2338..2356
 /note="Tn5 O end functional in transposition"
 /citations=[4]
 /rpt_type=inverted
 601 c 675 g 548 t
 BASE COUNT 532 a 601 c 675 g 548 t
 ORIGIN

Query Match 100.0%; Score 1001; DB 1; Length 2356;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGGAGGATCATCCAGCGCGGTCCCGGAAACGATTCCGAAGCCCAACCTTTCAATG 60
 DB 1790 CGCTGGAGGATCATCCAGCGCGGTCCCGGAAACGATTCCGAAGCCCAACCTTTCAATG 1731

QY 61 AAGCGCGGTGGATCGAATCTCGTGATGGCAGCTTGGGGTGGCTGGTGGCTGATT 120
 DB 1730 AAGCGCGGTGGATCGAATCTCGTGATGGCAGCTTGGGGTGGCTGGTGGCTGATT 1671

QY 121 TCGAACCCCAAGTCCCGCTCAGAAGAACTCGTCAAGAAAGCGGATAGAGCGATGCGCT 180
 DB 1670 TCGAACCCCAAGTCCCGCTCAGAAGAACTCGTCAAGAAAGCGGATAGAGCGATGCGCT 1611

QY 181 CGGAATCGGAGCGCGATACCGTAAAGCAGGAGGCGGTGAGCCATTCGCGCCAA 240
 DB 1610 CGGAATCGGAGCGCGATACCGTAAAGCAGGAGGCGGTGAGCCATTCGCGCCAA 1551

QY 241 GCTCTTCAGCAATATCAAGGTAGCCAAACCTATGCTCTGATAGCGGTCCGCCACACCA 300
 DB 1550 GCTCTTCAGCAATATCAAGGTAGCCAAACCTATGCTCTGATAGCGGTCCGCCACACCA 1491

QY 301 GCCGCCACAGTCGATGATCCAGAAAGCGGCATTTCCACCAATGATTCGCGAAGC 360
 DB 1490 GCCGCCACAGTCGATGATCCAGAAAGCGGCATTTCCACCAATGATTCGCGAAGC 1431

QY 361 AGGCATGCCATGGTTCAGCAGGATTCCTCGCGGTGGGCGATGCGCGCTTGAAGCTGG 420
 DB 1430 AGGCATGCCATGGTTCAGCAGGATTCCTCGCGGTGGGCGATGCGCGCTTGAAGCTGG 1371

QY 421 CGAACAGTTCGGTGGCGCGCGCCCTGATGCTCTTCGTCAGATCATCTGATCGACAA 480
 DB 1370 CGAACAGTTCGGTGGCGCGCGCCCTGATGCTCTTCGTCAGATCATCTGATCGACAA 1311

QY 481 GACCGGCTTCATCCGAGTACGCTCGCTCGATGCGATGTTTCGCTTGGTGGTGAATG 540
 DB 1310 GACCGGCTTCATCCGAGTACGCTCGCTCGATGCGATGTTTCGCTTGGTGGTGAATG 1251

QY 541 GCGAGTAGCCGGATCAAGCGTATGCGAGCCCGCGGANTTGATGCGCATGATGATGACTT 600
 DB 1250 GCGAGTAGCCGGATCAAGCGTATGCGAGCCCGCGGANTTGATGCGCATGATGATGACTT 1191

QY 601 TCTCGCGAGGAGCAAGGTGAGATGACAGGAGATCTGCCCCGGCACTTTCGCCCAATAGCA 660
 DB 1190 TCTCGCGAGGAGCAAGGTGAGATGACAGGAGATCTGCCCCGGCACTTTCGCCCAATAGCA 1131

QY 661 GCGAGTCCCTTCCCGCTTCACTGACACGTCGACAGCAGCTGCGGACGAGGACGCCCTCG 720
 DB 1130 GCGAGTCCCTTCCCGCTTCACTGACACGTCGACAGCAGCTGCGGACGAGGACGCCCTCG 1071

QY 721 TGCGCAGCCACGATAGCCGCTGCTGCTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTT 780
 DB 1070 TGCGCAGCCACGATAGCCGCTGCTGCTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTT 1011

QY 781 CGGTCTTGACAAAGAAACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 DB 1010 CGGTCTTGACAAAGAAACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 951

QY 841 AGCAGCCGATTTCTGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 DB 950 AGCAGCCGATTTCTGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 891

QY 901 GAGACCTCGTGCATCACTTGTTCATCATGCGAAGCAGTCTCATCTCTCTCTCTT 960
 DB 890 GAGACCTCGTGCATCACTTGTTCATCATGCGAAGCAGTCTCATCTCTCTCTCTT 831

QY 961 GATCAGATCTTATCCCTGCGCCATCAGATCTTGGCGGC 1001
 DB 830 GATCAGATCTTATCCCTGCGCCATCAGATCTTGGCGGC 790

RESULT 7

AY048743/c
 LOCUS AY048743 3267 bp DNA circular SYN 11-SEP-2001
 DEFINITION Template plasmid pKD4, complete sequence.
 ACCESSION AY048743
 VERSION AY048743.1 GI:15554332
 KEYWORDS
 SOURCE
 ORGANISM
 TEMPLATE plasmid pKD4
 ARTIFICIAL sequences; vectors.
 REFERENCE
 1 (bases 1 to 3267)
 Datsenko, K.A. and Wanner, B.L.
 One-step inactivation of chromosomal genes in Escherichia coli K-12
 using PCR products
 Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6640-6645 (2000)
 JOURNAL MEDLINE
 20300954
 PUBMED 10829079
 REFERENCE
 2 (bases 1 to 3267)
 Datsenko, K.A. and Wanner, B.L.
 Direct Substitution
 Submitted (29-JUL-2001) Biological Sciences, Purdue University,
 Lilly Hall of Life Sciences, West Lafayette, IN 47907, USA
 JOURNAL
 FEATURES
 Location/Qualifiers
 1. .3267
 /organism="Template plasmid pKD4"
 /mol_type="genomic DNA"
 /db_xref="taxon:170492"
 31..50
 /note="priming site 1"
 32..51
 /note="priming site 3"
 complement(51..98)
 /note="natural FRT site"
 459..1253
 /note="kanamycin resistance"
 /codon_start=1
 /transl_table=11
 /product="Tn5 neomycin phosphotransferase"
 /protein_id="AA02035.1"
 /db_xref="GI:15554333"

misc_feature
 /note="priming site 1"
 misc_feature
 /note="priming site 3"
 misc_feature
 complement(51..98)
 /note="natural FRT site"
 459..1253
 /note="kanamycin resistance"
 CDS
 /codon_start=1
 /transl_table=11
 /product="Tn5 neomycin phosphotransferase"
 /protein_id="AA02035.1"
 /db_xref="GI:15554333"

```

/misc_feature
1050..1109
/note="common priming site k1"
/note="common priming site k2"
1090..1109
/note="common priming site kt"
complement(1444..1489)
/note="FRT site"
1507..1526
/note="priming site 2"
1523..1509
/note="R6K gamma replication origin"
complement(1979..2839)
/note="bia; ampicillin resistance"
/codon_start=1
/transl_table=11
/product="beta-lactamase"
/protein_id="AAL02036.1"
/db_xref="GI:15554334"
/translation="MSIHFVRVALIPFAAFCFLPVFAHPETLVKVDADQLGARVGY
IELDLSGKILDSFPERPPNWSFKVLGAVLSRIDAGEOLGERLHYSDNIVLE
YSPVTKHLTDGTVRELCSAALTMSDNTAANLLTTIGGPKETATFLNNQDVLRL
DRPELNSAIPENDERTFPMVAMTTLKLLTGLSLTLASRQLLIQDMEADKVVAGL
LRSAIPAGNFIADKSGAGERSGRIIAALGPDGKPSRIVVIYTTGSAQNDERNRQIA
ETGASLIKHW"
2986..3245
/note="rnb transcription terminator"
BASE COUNT 793 a 798 c 869 g 807 t
ORIGIN

```

```

Query Match 100.0%; Score 1001; DB 12; Length 3267;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCTGGAGGATCATCAGCGGGCTCCCGGAAAACGATTCCGAAGCCCAACCTTTCATAG 60
Db 1392 CGCTGGAGGATCATCAGCGGGCTCCCGGAAAACGATTCCGAAGCCCAACCTTTCATAG 1333

Qy 61 AAGCGCGCGTGAATCGAAATCTCTGATGCGAGGTTGGGCGTCTGCTTGGTTCGCTATT 120
Db 1332 AAGCGCGCGTGAATCGAAATCTCTGATGCGAGGTTGGGCGTCTGCTTGGTTCGCTATT 1273

Qy 121 TCGAACCCAGATCCCGCTCAGAGAACTCTCAGAGAGCGGATAGAGCGGATGGCT 180
Db 1272 TCGAACCCAGATCCCGCTCAGAGAACTCTCAGAGAGCGGATAGAGCGGATGGCT 1213

Qy 181 CGGATCGGAGCGGCGATACCGTAAGCAAGAGAGCGGTCAGCCCATTCGCGGCAA 240
Db 1212 CGGATCGGAGCGGCGATACCGTAAGCAAGAGAGCGGTCAGCCCATTCGCGGCAA 1153

Qy 241 GCTCTTCAGCAATATCAGCGGTAGCAACGCTATGTCCTGATAGCGGTCGCGCAACCCA 300
Db 1152 GCTCTTCAGCAATATCAGCGGTAGCAACGCTATGTCCTGATAGCGGTCGCGCAACCCA 1093

Qy 301 GCGGCGCAGTCGATGAATCCAGAAAAGCGGCCATTTCACACCATGATATTCGCGAAGC 360
Db 1092 GCGGCGCAGTCGATGAATCCAGAAAAGCGGCCATTTCACACCATGATATTCGCGAAGC 1033

Qy 361 AGGCATCGCATGGTCAGCAGATCTCTCGCGTCGGGATCGCGCCCTTGAGCCCTGG 420
Db 1032 AGGCATCGCATGGTCAGCAGATCTCTCGCGTCGGGATCGCGCCCTTGAGCCCTGG 973

Qy 421 CGAACAGTTCGGTGGCGAGCCCTGATGCTCTTCGTCAGATCATCTCTGATCGACAA 480
Db 972 CGAACAGTTCGGTGGCGAGCCCTGATGCTCTTCGTCAGATCATCTCTGATCGACAA 913

Qy 481 GACCGGCTTCATCCAGTACGTCGCTCGATGCGATGTTTCGCTTGGTGGTCGATG 540

```

```

Db 912 GACCGGCTTCCATCCGAGTAGCTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTCAATG 853
Qy 541 GCGAGGTAGCCGAGTCAAGCGTATGCGAGCCGCGCATGTCATGAGCCATGATGATGACTT 600
Db 852 GCGAGGTAGCCGAGTCAAGCGTATGCGAGCCGCGCATGTCATGAGCCATGATGATGACTT 793
Qy 601 TCTCGCGAGGAGCAAGGTGAGATGACAGGAGATCCTGCCCGGCACTTCGCCCAATAGCA 660
Db 792 TCTCGCGAGGAGCAAGGTGAGATGACAGGAGATCCTGCCCGGCACTTCGCCCAATAGCA 733
Qy 661 GCGAGTCCCTTCCCGCTTCAGTGAACAGTGCAGACAGTGCAGAGAACGCGCCGTCG 720
Db 732 GCGAGTCCCTTCCCGCTTCAGTGAACAGTGCAGACAGTGCAGAGAACGCGCCGTCG 673
Qy 721 TGGCCAGCACGATAGCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 780
Db 672 TGGCCAGCACGATAGCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 613
Qy 781 CGGTCTTGACAAAAGAACCGCGGCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 840
Db 612 CGGTCTTGACAAAAGAACCGCGGCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 553
Qy 841 AGCAGCCGATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 900
Db 552 AGCAGCCGATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 493
Qy 901 GAGAACCTTGGTGCATTCATCTTGTTCATCATGCGAAAGCATCTCTCTGCTCTCTT 960
Db 492 GAGAACCTTGGTGCATTCATCTTGTTCATCATGCGAAAGCATCTCTCTGCTCTCTT 433
Qy 961 GATCAGATCTTGATCCCTCGCGCATCAGATCCTTGGCGGC 1001
Db 432 GATCAGATCTTGATCCCTCGCGCATCAGATCCTTGGCGGC 392

```

```

RESULT 8
A00782/c
LOCUS 3740 bp DNA linear PAT 16-FEB-1993
DEFINITION Plasmid pDM1.1 DNA sequence from patent EP0309746.
ACCESSION A00782
VERSION A00782.1 GI:14616
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences
REFERENCE 1 (bases 1 to 3740)
AUTHORS Certa,U.
TITLE Antimalaria vaccines
JOURNAL Patent: EP 0309746-A 11 05-APR-1989;
F. HOPMANN-LA ROCHE AG
FEATURES
location/Qualifiers
source
1..3740
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
BASE COUNT 813 a 1035 c 1056 g 836 t
ORIGIN
Query Match 100.0%; Score 1001; DB 6; Length 3740;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCTGGAGGATCATCAGCGGGCTCCCGGAAAACGATTCCGAAGCCCAACCTTTCATAG 60
Db 1290 CGCTGGAGGATCATCAGCGGGCTCCCGGAAAACGATTCCGAAGCCCAACCTTTCATAG 1231

Qy 61 AAGCGCGCGTGAATCGAAATCTCTGATGCGAGGTTGGGCGTCTGCTTGGTTCGCTATT 120
Db 1230 AAGCGCGCGTGAATCGAAATCTCTGATGCGAGGTTGGGCGTCTGCTTGGTTCGCTATT 1171

Qy 121 TCGAACCCAGATCCCGCTCAGAGAACTCTCAGAGAGCGGATAGAGCGGATGGCT 180
Db 1170 TCGAACCCAGATCCCGCTCAGAGAACTCTCAGAGAGCGGATAGAGCGGATGGCT 1111

```

```

QY 181 GCGAATCGGAGCGGCGATACCGTAAGACAGAGGAGCGGTACGCCCATTCGCGCCAA 240
Db 110 GCGAATCGGAGCGGCGGATACCGTAAGACAGAGGAGCGGTACGCCCATTCGCGCCAA 1051
QY 241 GCTCTTCAGCAATATCACGGGTAGCCAAAGCTATGTCCTGATAGCGGTCCGCCACACCA 300
Db 1050 GCTCTTCAGCAATATCACGGGTAGCCAAAGCTATGTCCTGATAGCGGTCCGCCACACCA 991
QY 301 GCGGCGCACAGTGGATGAAATCCAGAAAGCGGCCATTTCCACCATGATATTCGGCAAGC 360
Db 990 GCGGCGCACAGTGGATGAAATCCAGAAAGCGGCCATTTCCACCATGATATTCGGCAAGC 931
QY 361 AGGCATCGCCATGGGTACACAGAGATCTCTCGCGCTCGGCGCATGCGCGCTTGAACCTGG 420
Db 930 AGGCATCGCCATGGGTACACAGAGATCTCTCGCGCTCGGCGCATGCGCGCTTGAACCTGG 871
QY 421 CGACAGTTCGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
Db 870 CGACAGTTCGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 811
QY 481 GACCGGCTTCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTGGTGGT 540
Db 810 GACCGGCTTCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTGGTGGT 751
QY 541 GGCAGGTAGCGGATCAAGCGTATCAGCGCGCGCATTCGCATCAGCCATGATGATATCTT 600
Db 750 GGCAGGTAGCGGATCAAGCGTATCAGCGCGCGCATTCGCATCAGCCATGATGATATCTT 691
QY 601 TCTCGGAGGACAAAGTGGATGACAGAGATCTCTCGCGCGCATGCGCGCGCATGCGCGCG 660
Db 690 TCTCGGAGGACAAAGTGGATGACAGAGATCTCTCGCGCGCATGCGCGCGCATGCGCGCG 631
QY 661 GCGAGTTCGCTTCCGCTTCAGTGAACAGTGGAGCAGAGTGGCGCAAGGAGGAGCGCGCTG 720
Db 630 GCGAGTTCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTT 571
QY 721 TGCGCAGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
Db 570 TGCGCAGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 511
QY 781 CGGTCTTCGACAAAAGAGACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
Db 510 CGGTCTTCGACAAAAGAGACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 451
QY 841 AGCAGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
Db 450 AGCAGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 391
QY 901 GAGAACCTCGGTGGTGAATCCATCTTGTTCATCATGCGAAACGATCTCTATCTGTCCTT 960

```

RESULT 9

```

A02231/c 3740 bp DNA linear PAT 30-APR-1996
LOCUS A02231
DEFINITION Plasmid pDM1,1 DNA.
ACCESSION A02231
VERSION A02231.1 GI:490290
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences
REFERENCE 1 (bases 1 to 3740)
AUTHORS Doebeli,H., Eggmann,B., Gentz,R., Hochuli,B. and Stueber,D.
TITLE Fusion proteins and their purification
JOURNAL Patent: EP 0282042-A 5 14-SEP-1988;
F. HOFFMANN-LA ROCHE AG

```

FEATURES

Location/Qualifiers
 1..3740
 /organism="synthetic construct"
 /mol_type="genomic DNA"
 /db_xref="taxon:32630"

BASE COUNT 813 a 1035 c 1056 g 836 t
 ORIGIN

Query Match 100.0%; Score 1001; DB 6; Length 3740;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 1001; Conservative 0;

QY 1 CGCTGAGGATCATCAGCGCGCTCCCGGAAACGATTCGGAAGCCCAACCTTTTCATAG 60

Db 1290 CGCTGAGGATCATCAGCGCGCTCCCGGAAACGATTCGGAAGCCCAACCTTTTCATAG 1231

QY 61 AAGGCGGCGGTGAATCGAAATCTCGTATGCGAGGTGGCGTGGCGTGGCGTGGCGTGGCGT 120

Db 1230 AAGGCGGCGGTGAATCGAAATCTCGTATGCGAGGTGGCGTGGCGTGGCGTGGCGTGGCGT 1171

QY 121 TCGAACCACAGAGTCCCGCTCAGAGAACTCGTCAAGAAGCGGATAGAAGCGGATCGCT 180

Db 1170 TCGAACCACAGAGTCCCGCTCAGAGAACTCGTCAAGAAGCGGATAGAAGCGGATCGCT 1111

QY 181 GCGAATCGGAGCGGCGGATACCGTAAGACAGAGAGCGGTACGCCCATTCGCGCCAA 240

Db 1110 GCGAATCGGAGCGGCGGATACCGTAAGACAGAGAGCGGTACGCCCATTCGCGCCAA 1051

QY 241 GCTCTTCAGCAATATCACGGGTAGCCAAAGCTATGTCCTGATAGCGGTCCGCCACACCA 300

Db 1050 GCTCTTCAGCAATATCACGGGTAGCCAAAGCTATGTCCTGATAGCGGTCCGCCACACCA 991

QY 301 GCGGCGCACAGTGGATGAAATCCAGAAAGCGGCCATTTCCACCATGATATTCGGCAAGC 360

Db 990 GCGGCGCACAGTGGATGAAATCCAGAAAGCGGCCATTTCCACCATGATATTCGGCAAGC 931

QY 361 AGGCATCGCCATGGGTACACAGAGATCTCTCGCGCTCGGCGCATGCGCGCTTGAACCTGG 420

Db 930 AGGCATCGCCATGGGTACACAGAGATCTCTCGCGCTCGGCGCATGCGCGCTTGAACCTGG 871

QY 421 CGAAGTTCGCTTCCGCTTCAGTGAACAGTGGAGCAGAGTGGCGCAAGGAGGAGCGCGCTG 480

Db 870 CGAAGTTCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTT 811

QY 481 GACCGGCTTCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTGGTGGT 540

Db 810 GACCGGCTTCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTGGTGGT 751

QY 541 GGCAGGTAGCGGATCAAGCGTATCAGCGCGCATTCGCATCAGCCATGATGATATCTT 600

Db 750 GGCAGGTAGCGGATCAAGCGTATCAGCGCGCATTCGCATCAGCCATGATGATATCTT 691

QY 601 TCTCGGAGGACAAAGTGGATGACAGAGATCTCTCGCGCGCATGCGCGCGCATGCGCGCG 660

Db 690 TCTCGGAGGACAAAGTGGATGACAGAGATCTCTCGCGCGCATGCGCGCGCATGCGCGCG 631

QY 661 GCGAGTTCGCTTCCGCTTCAGTGAACAGTGGAGCAGAGTGGCGCAAGGAGGAGCGCGCTG 720

Db 630 GCGAGTTCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTT 571

QY 721 TGCGCAGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780

Db 570 TGCGCAGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 511

QY 781 CGGTCTTCGACAAAAGAGACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840

Db 510 CGGTCTTCGACAAAAGAGACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 451

QY 841 AGCAGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900

Db 450 AGCAGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 391

QY 901 GAGAACCTCGGTGGTGAATCCATCTTGTTCATCATGCGAAACGATCTCTATCTGTCCTT 960


```

Db 390 GAGAACCTCGGTGCAATCCATCTGTTCAATCATCGGAAACGATCCATCCCTGCTCTT 331
QY 961 GATCAGATCTTATCCCTGCGCATCAGATCCCTGCGGC 1001
Db 330 GATCAGATCTTATCCCTGCGCATCAGATCCCTGCGGC 290

RESULT 10
A02297/c
LOCUS A02297 3740 bp DNA linear PAT 18-MAY-1993
DEFINITION Plasmid pDML1 DNA.
ACCESSION A02297
VERSION A02297.1 GI:345260
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences
REFERENCE 1 (bases 1 to 3740)
AUTHORS Certa, U., Gentz, R. and Takacs, B.
TITLE Plasmodium falciparum merozoite antigen peptides
JOURNAL Patent: EP 0283829-A 18 28-SEP-1988;
F. HOFFMANN-LA ROCHE AG
FEATURES
Location/Qualifiers
1..3740
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
BASE COUNT 813 a 1035 c 1056 g 836 t
ORIGIN
Query Match 100.0%; Score 1001; DB 6; Length 3740;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGAGATCATCCAGCGCGTCCGGAACAGATTCGAAAGCCAACTTTTCATAG 60
Db 1290 CGCTGAGATCATCCAGCGCGTCCGGAACAGATTCGAAAGCCAACTTTTCATAG 1231
QY 61 AAGGGCGCGTGGATCGAATCTGTCATGCGAGGTTGGCGTCTGCTCGGTCAAT 120
Db 1230 AAGGGCGCGTGGATCGAATCTGTCATGCGAGGTTGGCGTCTGCTCGGTCAAT 1171
QY 121 TCGAACCCAGAGTCCCGCTCAGAAAGATCTGTCAGAAAGCGATAGAGCGGATCGCT 180
Db 1170 TCGAACCCAGAGTCCCGCTCAGAAAGATCTGTCAGAAAGCGATAGAGCGGATCGCT 1111
QY 181 GCGAATCGGGAGCGCGATACCGTAAAGCAGAGAGCGGTACGCCCATTCGCCGCCAA 240
Db 1110 GCGAATCGGGAGCGCGATACCGTAAAGCAGAGAGCGGTACGCCCATTCGCCGCCAA 1051
QY 241 GCTCTTCAGCAATATCAGGGTAGCCAAAGCTATGTCCTGATAGCGGTCCGCCACCCA 300
Db 1050 GCTCTTCAGCAATATCAGGGTAGCCAAAGCTATGTCCTGATAGCGGTCCGCCACCCA 991
QY 301 GCGGGCCAGTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATATTCGGAAGC 360
Db 990 GCGGGCCAGTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATATTCGGAAGC 931
QY 361 AGGCATCGCCATGCGGTACAGAGATCTCGCGGTGGGATCGCGGCTTTAGCGCTG 420
Db 930 AGGCATCGCCATGCGGTACAGAGATCTCGCGGTGGGATCGCGGCTTTAGCGCTG 871
QY 421 CGAACAGTTTCGCTCGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 480
Db 870 CGAACAGTTTCGCTCGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 811
QY 481 GACCGGCTTCCATCCGAGTACGTCCTCGCTCGATCGGATGTTTCGTTGGTTCGATG 540
Db 810 GACCGGCTTCCATCCGAGTACGTCCTCGCTCGATCGGATGTTTCGTTGGTTCGATG 751
QY 541 GCGAGTACCGGATCAAGCGGTATGCGCCCGCATTTGATCAGCGCATGATGATCACTT 600

```

```

Db 750 GCGAGTAGCCGATCAAGCGTATGAGCCGCCGATTCGATCGCATGATGATGACTT 891
QY 601 TCTCGGAGGAGCAAGGTGAGATGACAGGAGATCTGCCCCGGCACTTCGCCCAATAGCA 660
Db 690 TCTCGGAGGAGCAAGGTGAGATGACAGGAGATCTGCCCCGGCACTTCGCCCAATAGCA 631
QY 661 GCGAGTCCCTTCCCGCTTCAAGTCAACACGTGAGAGCAGCTGCGGCAAGAACCGCCGTCG 720
Db 630 GCGAGTCCCTTCCCGCTTCAAGTCAACACGTGAGAGCAGCTGCGGCAAGAACCGCCGTCG 571
QY 721 TGGCAGCAGCAGATAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 570 TGGCAGCAGCAGATAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 511
QY 781 CGGTCTTGACAAAAGAACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 510 CGGTCTTGACAAAAGAACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 451
QY 841 AGCAGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 450 AGCAGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 391
QY 901 GAGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 390 GAGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 331
QY 961 GATCAGATCTTATGATCCCTGCGCATCAGATCCCTTGGCGGC 1001
Db 330 GATCAGATCTTATGATCCCTGCGCATCAGATCCCTTGGCGGC 290

RESULT 11
A07054/c
LOCUS A07054 3740 bp DNA linear PAT 25-AUG-1993
DEFINITION Artificial sequence neo and lacI gene for neomycin
phosphotransferase and lac repressor.
ACCESSION A07054
VERSION A07054.1 GI:412971
KEYWORDS synthetic construct
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 3740)
AUTHORS Doebeli, H., Gentz, R. and Hochuli, E.
TITLE Homogeneous recombinant immune interferon fragments and
pharmaceutical compositions containing same
JOURNAL Patent: EP 0256424-A 11 24-FEB-1988;
F. HOFFMANN-LA ROCHE AG
FEATURES
Location/Qualifiers
1..3740
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
BASE COUNT 813 a 1035 c 1056 g 836 t
ORIGIN
Query Match 100.0%; Score 1001; DB 6; Length 3740;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGAGGATCATCCAGCGCGTCCGGAACAGATTCGAAAGCCAACTTTTCATAG 60
Db 1290 CGCTGAGGATCATCCAGCGCGTCCGGAACAGATTCGAAAGCCAACTTTTCATAG 1231
QY 61 AAGGGCGCGTGGATCGAATCTGTCATGCGAGGTTGGCGTCTGCTCGGTCAAT 120
Db 1230 AAGGGCGCGTGGATCGAATCTGTCATGCGAGGTTGGCGTCTGCTCGGTCAAT 1171
QY 121 TCGAACCCAGAGTCCCGCTCAGAAAGATCTGTCAGAAAGCGATAGAGCGGATCGCT 180
Db 1170 TCGAACCCAGAGTCCCGCTCAGAAAGATCTGTCAGAAAGCGATAGAGCGGATCGCT 1111
QY 181 GCGAATCGGGAGCGCGATACCGTAAAGCAGAGAGCGGTACGCCCATTCGCCGCCAA 240
Db 1110 GCGAATCGGGAGCGCGATACCGTAAAGCAGAGAGCGGTACGCCCATTCGCCGCCAA 1051
QY 241 GCTCTTCAGCAATATCAGGGTAGCCAAAGCTATGTCCTGATAGCGGTCCGCCACCCA 300
Db 1050 GCTCTTCAGCAATATCAGGGTAGCCAAAGCTATGTCCTGATAGCGGTCCGCCACCCA 991
QY 301 GCGGGCCAGTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATATTCGGAAGC 360
Db 990 GCGGGCCAGTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATATTCGGAAGC 931
QY 361 AGGCATCGCCATGCGGTACAGAGATCTCGCGGTGGGATCGCGGCTTTAGCGCTG 420
Db 930 AGGCATCGCCATGCGGTACAGAGATCTCGCGGTGGGATCGCGGCTTTAGCGCTG 871
QY 421 CGAACAGTTTCGCTCGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 480
Db 870 CGAACAGTTTCGCTCGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 811
QY 481 GACCGGCTTCCATCCGAGTACGTCCTCGCTCGATCGGATGTTTCGTTGGTTCGATG 540
Db 810 GACCGGCTTCCATCCGAGTACGTCCTCGCTCGATCGGATGTTTCGTTGGTTCGATG 751
QY 541 GCGAGTACCGGATCAAGCGGTATGCGCCCGCATTTGATCAGCGCATGATGATCACTT 600

```


Db 1110 GCGAATCGGAGCGCGGATACCGTAAAGCAGAGAAAGCGGTGAGCCCATTCGCGCCAA 1051
Qy 241 GCTCTTCAGCAATATACGGGTAGCCAAAGCTATGTCCTGATAGCGTCCGCCACACCCA 300
Db 1050 GCTCTTCAGCAATATACGGGTAGCCAAAGCTATGTCCTGATAGCGTCCGCCACACCCA 991
Qy 301 GCCGGCCACAGTCAGTCAATCCAGAAAGCGGCATTTTCCACATGATATTCGGAAGC 360
Db 990 GCCGGCCACAGTCAGTCAATCCAGAAAGCGGCATTTTCCACATGATATTCGGAAGC 931
Qy 361 AGGCATCGCCATGGGTACACAGAGATCTTCGCGCTCGGCATCGCGCTTCGAGCTGG 420
Db 930 AGGCATCGCCATGGGTACACAGAGATCTTCGCGCTCGGCATCGCGCTTCGAGCTGG 871
Qy 421 CGAACGTTTCGCTGGCGCGAGCCCTGATGCTCTTCGTCAGATCATCTGATCGACAA 480
Db 870 CGAACGTTTCGCTGGCGCGAGCCCTGATGCTCTTCGTCAGATCATCTGATCGACAA 811
Qy 481 GACCGGCTTCATCCGAGTACGTCTCGCTCGATGCGATGTTTCGCTTCGTCAGATCATCTGATCGACAA 840
Db 810 GACCGGCTTCATCCGAGTACGTCTCGCTCGATGCGATGTTTCGCTTCGTCAGATCATCTGATCGACAA 811
Qy 541 GGCAGTAGCCGATCAAGCTATGAGCGCGCGCATTTGATCAGCCATGATGATGAT 600
Db 750 GGCAGTAGCCGATCAAGCTATGAGCGCGCGCATTTGATCAGCCATGATGATGAT 591
Qy 601 TCTCGGAGCAGCAAGGTGATGACAGAGATCTTCGCGCTCGGCATCGGCATGATGATGAT 660
Db 590 TCTCGGAGCAGCAAGGTGATGACAGAGATCTTCGCGCTCGGCATCGGCATGATGATGAT 631
Qy 661 GCGAGTCCCTTCGCTTCAGTGACAAAGCTCGAGCAGCTCGCGAGGAAAGCGCGCTG 720
Db 630 GCGAGTCCCTTCGCTTCAGTGACAAAGCTCGAGCAGCTCGCGAGGAAAGCGCGCTG 571
Qy 721 TGSCCAGCAGTACGCGCTGCTCTGCTGAGTTCATTCAGGCAACCGGACAGGT 780
Db 570 TGSCCAGCAGTACGCGCTGCTCTGCTGAGTTCATTCAGGCAACCGGACAGGT 511
Qy 781 CGGTCTTGACAAAAGAACCGCGCGCTGACAGCGGAAACAGCGCGGATCAG 840
Db 510 CGGTCTTGACAAAAGAACCGCGCGCTGACAGCGGAAACAGCGCGGATCAG 451
Qy 841 AGCAGCGGATGCTGTTGTGCGCAGTCAATAGCCGAATAGCTCTCCACCCAGCGCG 900
Db 450 AGCAGCGGATGCTGTTGTGCGCAGTCAATAGCCGAATAGCTCTCCACCCAGCGCG 391
Qy 901 GAGAACCTGCGTCAATCCATCTGTTCAATCATCGGAAAGCATCTCATCTGCTCTT 960
Db 390 GAGAACCTGCGTCAATCCATCTGTTCAATCATCGGAAAGCATCTCATCTGCTCTT 331

RESULT 12

A10362/c

LOCUS

DEFINITION

Accession

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

A10362 Artificial sequence of plasmid pDM1.1. linear PAT 22-SEP-1993

A10362 Artificial sequence of plasmid pDM1.1. linear PAT 22-SEP-1993

A10362.1 GI:490684

synthetic construct

artificial sequences

1 (bases 1 to 3740)

Bujard,H. and Lanzer,M.

High repressible sequence for control of expression

Patent: EP 0303925-A 9 22-FEB-1989;

F. HOFFMANN-LA ROCHE AG

Location/Qualifiers

1..3740

```

QY 961 GATCAGATCTTGATCCCTCGCCATCAGATCCTTGGCGGC 1001
Db 330 GATCAGATCTTGATCCCTCGCCATCAGATCCTTGGCGGC 290

RESULT 13
AL14595/c
LOCUS Al14595
DEFINITION Synthetic nucleotide sequence of the pDm1,1 plasmid.
ACCESSION Al14595
VERSION Al14595.1
KEYWORDS GI:491836
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 3740)
AUTHORS Bannwarth, W., Certa, U., Mous, J., and Stueber, D.
TITLE Polypeptides that elicit antibodies against AIDS virus
JOURNAL Patent: EP 0219106-A 8 22-APR-1987;
F. HOFFMANN-LA ROCHE AG
FEATURES
    source
        Location/Qualifiers
            1..3740
                /organism="synthetic construct"
                /mol_type="genomic DNA"
                /db_xref="taxon:32630"
                357..1151
                    /codon_start=1
                    /transl_table=1
                    /product="neomycin phosphotransferase"
                    /protein_id="CAA01174.1"
                    /db_xref="GI:491837"
                    /translation="MIEQDLHAGSPAWVRELFGVDWAQQTIGSDAAVFLSNAQGR
                    PVLVKTDLSGALNELDEARLSHLATGTGVCFAVLVVTEAGRDWLLGEVFGQDL
                    LSHLAPAEKVSIMADAMRRRLHTLPATCFPDQAKHRIERARTMERAGLVQDDLDE
                    EQGLAPAEKVSIMADAMRRRLHTLPATCFPDQAKHRIERARTMERAGLVQDDLDE
                    OIALATRDIAEELGGEWADRFVLVYGLAAPDSORIAFYRLIDEPF"
                    complement(1529..2611)
                    /codon_start=1
                    /transl_table=1
                    /product="lac repressor"
                    /protein_id="CAA01175.1"
                    /db_xref="GI:4526670"
                    /translation="MKPVTLYDAEYAGVSYQTVSRVVNQASHVSAKTREKVEAAMAE
                    LNYIPNRVAQQLAGKSLIIVATSLALHAPSQIVAAIKSRADQLGASVVSVMSERS
                    GVEACAAVNLHIAORVSLIINVLDDQDAIAVEAACTNVPALPLDVSQDTPLNSII
                    FSHEDGTRLVGHVYALGHQIALLAGLPSVSAIRLHAGHKILTRNQIPIAREG
                    DRSMSGFOOTQMNLGIVPTAMLVANDQWALGAMRAITSEGLRVGADISVWGVDIT
                    EDSSCTIPPTTIKQDFRLGQTSVDRLLQLSQQAQVKGNNQLLPVSLVKKRTTLAPNT
                    QPASPRALADSLMLQARQVSRLESQ"
BASE COUNT 813 a 1036 c 1055 g 836 t
ORIGIN

Query Match 100.0%; Score 1001; DB 6; Length 3740;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGAGGATCATCCAGCGCGCTCCCGGAAAACGATCCGAAGCCCAACCTTTCATAG 60
Db 1290 CGCTGAGGATCATCCAGCGCGCTCCCGGAAAACGATCCGAAGCCCAACCTTTCATAG 1231

QY 61 AAGCGCGCGTGGATCGAAATCTCGTATGCGAGGTGGCGTGGCGTGGCGTGGCGTCAAT 120
Db 1230 AAGCGCGCGTGGATCGAAATCTCGTATGCGAGGTGGCGTGGCGTGGCGTGGCGTCAAT 1171

QY 121 TCGAACCCCGAGTCCCGCTCAGAAAGACTCTCGTAAAGGCGGATAGAGCGGATGCGCT 180
Db 1170 TCGAACCCCGAGTCCCGCTCAGAAAGACTCTCGTAAAGGCGGATAGAGCGGATGCGCT 1111

QY 181 GCGAATCGGAGCGCGGATACCGTAAAGCAGGAGCGGTGACGCCCATTCGCGCGCAA 240
Db 1110 GCGAATCGGAGCGCGGATACCGTAAAGCAGGAGCGGTGACGCCCATTCGCGCGCAA 1051

LOCUS Al14595/c
DEFINITION Artificial sequence of plasmid pDm1,1.
ACCESSION Al14595
VERSION Al14595.1
KEYWORDS GI:513600
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 3740)
AUTHORS Bannwarth, W., Caspers, P., Le Grice, S., and Mous, J.
TITLE Recombinant HIV-2 polypeptides
JOURNAL Patent: EP 0316695-A 6 24-MAY-1989;
F. HOFFMANN-LA ROCHE AG
FEATURES
    source
        Location/Qualifiers
            1..3740
                /organism="synthetic construct"
                /mol_type="genomic DNA"
                /db_xref="taxon:32630"

```

BASE COUNT 813 a 1035 c 1056 g 836 t

Query Match 100.0%; Score 1001; DB 6; Length 3740;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGGAGATCATCCAGCGCGGTCCGGAAGACGATTCGGAAGCCCAACCTTTCATAG 60
DB 1290 CGCTGGAGATCATCCAGCGCGGTCCGGAAGACGATTCGGAAGCCCAACCTTTCATAG 1231

QY 61 AAGCGCGGTGGGAATCGAATCGATGCGAGGTGGCGTCGCTTGGTGGTCAAT 120
DB 1230 AAGCGCGGTGGGAATCGAATCGATGCGAGGTGGCGTCGCTTGGTGGTCAAT 1171

QY 121 TCGAACCCAGAGTCCCGCTCAGAGAATCGTCAAGAAGCGGATAGAGCGGATCGCT 180
DB 1170 TCGAACCCAGAGTCCCGCTCAGAGAATCGTCAAGAAGCGGATAGAGCGGATCGCT 1111

QY 181 GCGAATCGGAGCGCGATACCGTAAGACGAGGAGCGGTGAGCCCATTCGCGCCAA 240
DB 1110 GCGAATCGGAGCGCGATACCGTAAGACGAGGAGCGGTGAGCCCATTCGCGCCAA 1051

QY 241 GCTCTTCAGCAATATCAGCGGTAGCCAAACGCTATGCTCGATAGCGGTCCGCCACCCCA 300
DB 1050 GCTCTTCAGCAATATCAGCGGTAGCCAAACGCTATGCTCGATAGCGGTCCGCCACCCCA 991

QY 301 GCCGGCCACATGATGATCCAGAAAGCGGATATTCACCATGATATTCGGCAAGC 360
DB 990 GCCGGCCACATGATGATCCAGAAAGCGGATATTCACCATGATATTCGGCAAGC 931

QY 361 AGGCATCGCCATGAGTCCCGCTCAGAGAATCGTCAAGAAGCGGATAGAGCGGATCGCT 420
DB 930 AGGCATCGCCATGAGTCCCGCTCAGAGAATCGTCAAGAAGCGGATAGAGCGGATCGCT 871

QY 421 CGAACAGTTCGGCTGGCGGAGCGGCTGATGCTTTCGTCAGATCATTCGATCGCAAA 480
DB 870 CGAACAGTTCGGCTGGCGGAGCGGCTGATGCTTTCGTCAGATCATTCGATCGCAAA 811

QY 481 GACCGGTTCATCCAGATGATGCTCGCTCGATGCGATGTTTCGCTTGGTGGTCAATG 540
DB 810 GACCGGTTCATCCAGATGATGCTCGCTCGATGCGATGTTTCGCTTGGTGGTCAATG 751

QY 541 GGCAGGTAGCGGATCAAGCGTATGACGCCCGCATTCGATCAGCCATGATGATCACT 600
DB 750 GGCAGGTAGCGGATCAAGCGTATGACGCCCGCATTCGATCAGCCATGATGATCACT 591

QY 601 TCTCGGAGGACGAGGTGAGTACAGAGGATGCTCGCCCGGACATTCGCCCAATAGCA 660
DB 590 TCTCGGAGGACGAGGTGAGTACAGAGGATGCTCGCCCGGACATTCGCCCAATAGCA 631

QY 661 GATCAGATCTTTGATCCCGCTCGCCATCAGATCTTGGCGGC 1001

DB 330 GATCAGATCTTCATCCCGCTCGCCATCAGATCTTGGCGGC 290

RESULT 15
A30452/c

LOCUS A30452 3740 bp DNA linear PAT 05-JUL-2002
DEFINITION plasmid pDM1.1 sequence.
ACCESSION A30452
VERSION A30452.1 GI:23957104
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial construct
REFERENCE 1 (bases 1 to 3740)
AUTHORS Fountoulakis M., Garotta, G. and Stueber, D.
TITLE Soluble interferon-gamma receptors and methods for their production
JOURNAL Patent: EP 0393502-A 16 24-OCT-1990;
F. HOFFMANN-LA ROCHE AG

FEATURES
Location/Qualifiers
1..3740
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

BASE COUNT 813 a 1035 c 1056 g 836 t
ORIGIN

Query Match 100.0%; Score 1001; DB 6; Length 3740;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGGAGATCATCCAGCGCGGTCCGGAAGACGATTCGGAAGCCCAACCTTTCATAG 60
DB 1290 CGCTGGAGATCATCCAGCGCGGTCCGGAAGACGATTCGGAAGCCCAACCTTTCATAG 1231

QY 61 AAGCGCGGTGGGAATCGAATCGATGCGAGGTGGCGTCGCTTGGTGGTCAAT 120
DB 1230 AAGCGCGGTGGGAATCGAATCGATGCGAGGTGGCGTCGCTTGGTGGTCAAT 1171

QY 121 TCGAACCCAGAGTCCCGCTCAGAGAATCGTCAAGAAGCGGATAGAGCGGATCGCT 180
DB 1170 TCGAACCCAGAGTCCCGCTCAGAGAATCGTCAAGAAGCGGATAGAGCGGATCGCT 1111

QY 181 GCGAATCGGAGCGCGATACCGTAAGACGAGGAGCGGTGAGCCCATTCGCGCCAA 240
DB 1110 GCGAATCGGAGCGCGATACCGTAAGACGAGGAGCGGTGAGCCCATTCGCGCCAA 1051

QY 241 GCTCTTCAGCAATATCAGCGGTAGCCAAACGCTATGCTCGATAGCGGTCCGCCACCCCA 300
DB 1050 GCTCTTCAGCAATATCAGCGGTAGCCAAACGCTATGCTCGATAGCGGTCCGCCACCCCA 991

QY 301 GCCGGCCACATGATGATCCAGAAAGCGGATATTCACCATGATATTCGGCAAGC 360
DB 990 GCCGGCCACATGATGATCCAGAAAGCGGATATTCACCATGATATTCGGCAAGC 931

QY 361 AGGCATCGCCATGAGTCCCGCTCAGAGAATCGTCAAGAAGCGGATAGAGCGGATCGCT 420
DB 930 AGGCATCGCCATGAGTCCCGCTCAGAGAATCGTCAAGAAGCGGATAGAGCGGATCGCT 871

QY 421 CGAACAGTTCGGCTGGCGGAGCGGCTGATGCTTTCGTCAGATCATTCGATCGCAAA 480
DB 870 CGAACAGTTCGGCTGGCGGAGCGGCTGATGCTTTCGTCAGATCATTCGATCGCAAA 811

QY 481 GACCGGTTCATCCAGATGATGCTCGCTCGATGCGATGTTTCGCTTGGTGGTCAATG 540
DB 810 GACCGGTTCATCCAGATGATGCTCGCTCGATGCGATGTTTCGCTTGGTGGTCAATG 751

QY 541 GGCAGGTAGCGGATCAAGCGTATGACGCCCGCATTCGATCAGCCATGATGATCACT 600
DB 750 GGCAGGTAGCGGATCAAGCGTATGACGCCCGCATTCGATCAGCCATGATGATCACT 591

QY 601 TCTCGGAGGACGAGGTGAGTACAGAGGATGCTCGCCCGGACATTCGCCCAATAGCA 660
DB 590 TCTCGGAGGACGAGGTGAGTACAGAGGATGCTCGCCCGGACATTCGCCCAATAGCA 631

```
Qy 661 GCCAGTCCCTTCCCGCTTCAGTGACAAACGTCCAGCACAGCTGCGCAAGGAACGCCCGTCG 720
Db |||||
Qy 630 GCCAGTCCCTTCCCGCTTCAGTGACAAACGTCCAGCACAGCTGCGCAAGGAACGCCCGTCG 571
Db |||||
Qy 721 TGGCCAGCCACGATAGCCGCGCTGCTCGTCTGTCAGTTTCATTGAGGGCACCGGACAGGT 780
Db |||||
Qy 570 TGGCCAGCCACGATAGCCGCGCTGCTCGTCTGTCAGTTTCATTGAGGGCACCGGACAGGT 511
Db |||||
Qy 781 CGGTCTTGACAAAAGAACCGGGCGCCCTCGCTGACAGCGGGAACACGGGGCATCAG 840
Db |||||
Qy 510 CGGTCTTGACAAAAGAACCGGGCGCCCTCGCTGACAGCGGGAACACGGGGCATCAG 451
Db |||||
Qy 841 AGCAGCGGATGTCTGTGTGCCAGTCATAGCCGAATAGCCTCTCCACCCGAAGCGGCG 900
Db |||||
Qy 450 AGCAGCGGATGTCTGTGTGCCAGTCATAGCCGAATAGCCTCTCCACCCGAAGCGGCG 391
Db |||||
Qy 901 GAGAACTGCGTGCAATCCATCTTGTTCATCATGCGAAACGATCCTCATCCTGTCTCTT 960
Db |||||
Qy 390 GAGAACTGCGTGCAATCCATCTTGTTCATCATGCGAAACGATCCTCATCCTGTCTCTT 331
Db |||||
Qy 961 GATCAGATCTTGATCCCTGCGCCATCAGATCCTTGGCGGC 1001
Db |||||
Qy 330 GATCAGATCTTGATCCCTGCGCCATCAGATCCTTGGCGGC 290
Db |||||
```

Search completed: February 16, 2004, 01:25:58
Job time : 3968.99 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 19:07:11 ; Search time 324.065 Seconds
 (without alignments)
 8338.264 Million cell updates/sec

Title: US-09-921-143-36_COPY_4000_5000
 Perfect score: 1001
 Sequence: 1 cgcggagatcatccagcc.....gcatcagatccttgaggcc 1001

Scoring table: OLIGO_NUC
 Gapop 60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq 19Jun03.*

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
 23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
 24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
 25: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	
1	1001	100.0	2220	21	AAZ58064	
C	2	1001	100.0	3176	14	AAQ42402
C	3	1001	100.0	3176	20	AAQ01440
C	4	1001	100.0	3177	13	AAQ27504
C	5	1001	100.0	3740	9	AAH80959
C	6	1001	100.0	3740	9	AAH81154
C	7	1001	100.0	3740	9	AAH80483
C	8	1001	100.0	3740	11	AAQ06307
						Vector plasmid pUN
						Tn5seq1 transposon
						Tn5seq1 transposon
						Transposon Tn5seq1
						Plasmid pDM1.1 enc
						Plasmid pDM1.1. P
						pDM1.1 contg. the
						Sequence of plasm

C	9	1001	100.0	3740	12	AAQ12786
C	10	1001	100.0	3740	14	AAQ34610
	11	1001	100.0	3974	19	AAV33721
	12	1001	100.0	3974	19	AAV66372
	13	1001	100.0	3974	19	AAV66999
	14	1001	100.0	3974	19	AAV57971
	15	1001	100.0	3974	19	AAV42677
	16	1001	100.0	3974	20	AAZ32002
	17	1001	100.0	3974	20	AAZ19915
	18	1001	100.0	3974	20	AAZ10532
	19	1001	100.0	3974	20	AAZ11171
	20	1001	100.0	3974	20	AAZ87252
	21	1001	100.0	3974	20	AAZ18994
	22	1001	100.0	3974	20	AAZ08956
	23	1001	100.0	3974	21	AAZ75540
	24	1001	100.0	3974	21	AAZ72757
	25	1001	100.0	3974	21	AAZ48490
	26	1001	100.0	3974	21	AAZ71310
	27	1001	100.0	3974	21	AAZ27353
	28	1001	100.0	3974	22	AAZ32008
	29	1001	100.0	3974	22	AAZ00059
	30	1001	100.0	3974	24	AAZ54432
	31	1001	100.0	3974	24	AAZ73147
	32	1001	100.0	3974	24	AAZ10061
	33	1001	100.0	3974	24	AAZ23719
	34	1001	100.0	3974	25	AAZ17117
	35	1001	100.0	3974	25	AAZ76975
	36	1001	100.0	3974	25	AAZ49531
	37	1001	100.0	3974	25	AAZ57778
	38	1001	100.0	3974	25	AAZ83101
	39	1001	100.0	3984	25	AAZ33990
	40	1001	100.0	3997	24	AAZ78682
	41	1001	100.0	4208	19	AAZ4537
	42	1001	100.0	4208	20	AAZ33993
	43	1001	100.0	4208	22	AAZ82961
	44	1001	100.0	4256	19	AAZ63305
	45	1001	100.0	4256	22	AAZ21539

ALIGNMENTS

RESULT 1
 AAZ58064
 ID AAZ58064 standard; DNA; 2220 BP.

XX AC AAZ58064;

DT 25-APR-2000 (first entry)

XX Vector plasmid pUNI-10.

DE Plasmid pUNI-10; vector; site-specific recombination; subcloning;
 Univector Fusion System; Univector Plasmid-Fusion System; ss.

KW Chimeric - Escherichia coli.

OS Chimeric - Bos taurus.

OS Chimeric - Bacteriophage T7

OS Chimeric - Synthetic.

XX Key Location/Qualifiers

FT rep_origin 1..400

FT /*tag= a

FT /*note= "R6K-gamma origin of replication"

FT misc_feature 401..414

FT /*tag= b

FT /*note= "polylinker facilitating lox site exchange"

FT misc_feature 415..448

FT /*tag= c

FT /*note= "wild-type loxP recombinase target site"

FT misc_feature 449..527

FT /*tag= d

FT /*note= "polylinker used as DNA insertion site"

pDMI.1. Unidentif
 plasmid pDMI.1. S
 Bacterial expressi
 Vector pHE4a. Chi
 plasmid pHE4-5 con
 vector pHE4-5. Ch
 pHEA-5 expression
 Vector pHE4-5. Ch
 Nucleotide sequenc
 pHE4-5 vector, con
 vector plasmid pHE
 PAFAI pHE4-5 expre
 Expression vector
 Nucleotide sequenc
 Expression vector
 pHE4a vector nucle
 pHE4-5 expression
 pHE4A expression v
 pHE4-5 expression
 pHE4-5 expression
 Human AIM II relat
 Expression vector
 pHE4-5 expression
 Human VEGF-2 relat
 vector pHE4a DNA S
 pHE4a expression
 pHE4-5 expression
 Expression vector
 Nucleotide sequenc
 pHE4-5 vector geno
 Expression vector
 Nucleotide sequenc
 Nucleotide sequenc
 pHE4-5 vector sequ

FT polyA_site 528..750
 FT /*tag= e
 FT /note= "BGH polyA site"
 FT terminator 751..890
 FT /*tag= f
 FT /note= "T7 terminator"
 FT prim_transcript 896..2220
 FT /*tag= g
 FT /note= "kanamycin resistance gene from Tn5"
 XX WO200005355-A1.
 XX PN
 XX PD 03-FEB-2000.
 XX PF 27-JUL-1998; 98WO-US15481.
 XX PR 24-JUL-1998; 98US-0122384.
 XX PA (BAYU) BAYLOR COLLEGE MEDICINE.
 XX PI Elledge SJ, Liu Q;
 XX DR WPI; 2000-171429/15.
 XX
 CC New methods for recombination of nucleic acid constructs, useful for
 CC the rapid subcloning of nucleic acid sequences in vivo and in vitro
 CC
 CC Claim 44; Page 63-64; 110pp; English.
 CC
 CC The present sequence is that of novel vector plasmid pUNI-10. It
 CC includes: the conditional origin of replication from RepK-gamma, that
 CC derived from plasmid RepK (ATCC 37120); a NotI-XpnI polylinker that
 CC facilitates the exchange of lox sites; a wild-type loxP site (see
 CC also AA258066); a polylinker used for the insertion of genomic DNA
 CC or cDNA sequences; a bovine growth hormone polyA site; a T7
 CC terminator; and a kanamycin resistance selectable marker gene from
 CC transposon Tn5, modified to removed the native NcoI site. pUNI-10
 CC is an example of novel Univectors or pUNI vectors of the invention
 CC that have a sequence-specific recombinase target site (e.g. loxP)
 CC preceding the insertion site for the gene of interest, a selectable
 CC marker gene (optional) and a conditional origin of replication that
 CC is active only in host cells expressing the requisite transacting
 CC replication factor (optional). The vectors are designed to contain
 CC a gene of interest but to lack a promoter for expression of the
 CC gene. The vectors are used in a novel method for the rapid
 CC subcloning of nucleic acid sequences in vivo and in vitro without
 CC the need of restriction endonucleases. The method is referred to
 CC as the Univector Fusion System or Univector Plasmid-fusion system
 CC (UPS). The UPS uses site-specific recombination to catalyze plasmid
 CC fusion between a Univector and host vectors containing regulatory
 CC information. In some embodiments, plasmid fusion events are
 CC genetically selected and result in placement of the gene of interest
 CC under the control of novel regulatory elements. A second UPS-related
 CC method allows for the precise transfer of coding sequences alone from
 CC a Univector into a host vector. UPS further provides means for the
 CC subcloning of entire nucleic acid libraries and the directional
 CC cloning of linear nucleic acid molecules, e.g. PCR products.
 XX
 XX Sequence 2220 BP; 506 A; 610 C; 581 G; 523 T; 0 other;
 XX
 XX Query Match 100.0%; Score 1001; DB 21; Length 2220;
 XX Best Local Similarity 100.0%; Pred. No. 0;
 XX Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGCTGGAGGATCATCCAGCGCGCTCCCGGAAAGATTCGGAAGCCCAACCTTTCATAG 60
 Db 932 CGCTGGAGGATCATCCAGCGCGCTCCCGGAAAGATTCGGAAGCCCAACCTTTCATAG 991
 QY 61 AAGCGCGCGTGGAAATCGAAATCTCGTATGCGAGGTTGGCGCTCGCTTGGTGCATTT 120
 Db 992 AAGCGCGCGTGGAAATCGAAATCTCGTATGCGAGGTTGGCGCTCGCTTGGTGCATTT 1051
 QY 121 TCGAACCCTCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGCGATAGAGGCGATCGCT 180

Db 1052 TCGAACCCTCAGAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGCGCGCT 1111
 QY 181 GCGAATCGGAGCGCGGATACCGTAAAGCAAGAGGAGCGGTGAGCCATTCGCGGCAA 240
 Db 1112 GCGAATCGGAGCGCGGATACCGTAAAGCAAGAGGAGCGGTGAGCCATTCGCGGCAA 1171
 QY 241 GCTCTTCAGCAATATACAGCGGTAGCCAAAGCTATGCTCTGATAGCGGTCCGCGACACCA 300
 Db 1172 GCTCTTCAGCAATATACAGCGGTAGCCAAAGCTATGCTCTGATAGCGGTCCGCGACACCA 1231
 QY 301 GCGGCGCACAGTGTGATGAATCCAGAAAGAGGCGCAATTTCCACATGATATTCGGAAGC 360
 Db 1232 GCGGCGCACAGTGTGATGAATCCAGAAAGAGGCGCAATTTCCACATGATATTCGGAAGC 1291
 QY 361 AGGCATCGCATGGTCAAGACGAGATCTTCGCGCTCGGCGCATCGCGCTTCAGGCTGG 420
 Db 1292 AGGCATCGCATGGTCAAGACGAGATCTTCGCGCTCGGCGCATCGCGCTTCAGGCTGG 1351
 QY 421 CGAACAGTTCGGTGGCGGAGCCCTCGATGCTCTTCGTCCAGATCATCTGATGACAA 480
 Db 1352 CGAACAGTTCGGTGGCGGAGCCCTCGATGCTCTTCGTCCAGATCATCTGATGACAA 1411
 QY 481 GACCGGCTTCATCCGAGTACGCTCGCTCGATGCGATGTTTCGCTTGGTGTGCAATG 540
 Db 1412 GACCGGCTTCATCCGAGTACGCTCGCTCGATGCGATGTTTCGCTTGGTGTGCAATG 1471
 QY 541 GGCAGGTAGCCGGATCAAGCGTATGCAGCGCGCCGCAATTCGATCAGCGCATGATCGATCTT 600
 Db 1472 GGCAGGTAGCCGGATCAAGCGTATGCAGCGCGCCGCAATTCGATCAGCGCATGATCGATCTT 1531
 QY 601 TCTTCGCGAGGACAAAGTGTGATGACAGGAGATCTTCGCGCGGCACTTCGCCCAATAGCA 660
 Db 1532 TCTTCGCGAGGACAAAGTGTGATGACAGGAGATCTTCGCGCGGCACTTCGCCCAATAGCA 1591
 QY 661 GCCAGTCCCTTCGCGCTTCAGTGACAAAGTTCGAGCACAGCTGGCGCAAGAACGCCGCTCG 720
 Db 1592 GCCAGTCCCTTCGCGCTTCAGTGACAAAGTTCGAGCACAGCTGGCGCAAGAACGCCGCTCG 1651
 QY 721 TGGCGAGCGACGATAGCGCGCTCGCTCGTCTGCGAGTTCATTCAGGGGACCGGACAGGT 780
 Db 1652 TGGCGAGCGACGATAGCGCGCTCGCTCGTCTGCGAGTTCATTCAGGGGACCGGACAGGT 1711
 QY 781 CGGTCTTGACAAAAGAACCGCGCGCCCTCGCTCGACAGCGGCAACACGCGGCGCATCAG 840
 Db 1712 CGGTCTTGACAAAAGAACCGCGCGCCCTCGCTCGACAGCGGCAACACGCGGCGCATCAG 1771
 QY 841 AGCAGCGGATGTCTGTGTGCGCCAGTTCATAGCCGAATAGCCTTCACCCCAAGCGCGCG 900
 Db 1772 AGCAGCGGATGTCTGTGTGCGCCAGTTCATAGCCGAATAGCCTTCACCCCAAGCGCGCG 1831
 QY 901 GAGAACCTGCGTCAATCCATCTTGTTCATCATGCGAAAGCATCTCATCTCTGTCTTT 960
 Db 1832 GAGAACCTGCGTCAATCCATCTTGTTCATCATGCGAAAGCATCTCATCTCTGTCTTT 1891
 QY 961 GATCAGATCTTGATCCCTTCGCGCATCAGATCCCTTCGCGCGC 1001
 Db 1892 GATCAGATCTTGATCCCTTCGCGCATCAGATCCCTTCGCGCGC 1932

RESULT 2
 AAQ42402/c
 ID AAQ42402 standard; DNA; 3176 BP.
 XX
 XX AAQ42402;
 XX AC
 XX DT 25-MAR-2003 (updated)
 XX DT 15-SEP-1993 (first entry)
 XX XX
 XX Tn5seq1 transposon sequence.
 KW Bidirectional sequencing; restriction enzyme cleavage sites; marker;
 KW Tn5 transposase gene; E. coli; ss.

[illegible]

QY 361 AGGATCCGATCGGTCACACAGATCCCTCGCGTCGGGCATGCGGCTTGAGCCTGG 420
 Db 980 AGGATCCGATCGGTCACACAGATCCCTCGCGTCGGGCATGCGGCTTGAGCCTGG 921
 QY 421 CGAACAGTTCGGTGGCGGAGCCCTGATGCTTCCTCCAGATCATCTGATGACAA 480
 Db 920 CGAACAGTTCGGTGGCGGAGCCCTGATGCTTCCTCCAGATCATCTGATGACAA 861
 QY 481 GACCGGTTCCATCCGAGTACGTGCTGCTCGATGCGATGTTTCGGTTCGATG 540
 Db 860 GACCGGTTCCATCCGAGTACGTGCTGCTCGATGCGATGTTTCGGTTCGATG 801
 QY 541 GGCAGGTAGCCGATCAAGGTATGACGCGCGCATTCATCAGCCATGATGATCTT 600
 Db 800 GGCAGGTAGCCGATCAAGGTATGACGCGCGCATTCATCAGCCATGATGATCTT 741
 QY 601 TCTCGGAGGAGCAAGGTAGATGACAGGAGATCTTGCCTCCGCACTTCGCCAATAGCA 660
 Db 740 TCTCGGAGGAGCAAGGTAGATGACAGGAGATCTTGCCTCCGCACTTCGCCAATAGCA 681
 QY 661 GCCAGTCCCTCCGCTTCAGTGACACAGTGCAGCAGTGGCGAAGAAAGCCGCTG 720
 Db 680 GCCAGTCCCTCCGCTTCAGTGACACAGTGCAGCAGTGGCGAAGAAAGCCGCTG 621
 QY 721 TGCCAGCCAGATAGCCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 Db 620 TGCCAGCCAGATAGCCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 561
 QY 781 CGGCTTCGACAAAGAACCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 Db 560 CGGCTTCGACAAAGAACCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 501
 QY 841 AGCAGCGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 Db 500 AGCAGCGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 441
 QY 901 GAGAACTGCTGATTCATCTGTTTCAATGCGAAACGATGCTGCTGCTGCTGCTGCT 960
 Db 440 GAGAACTGCTGATTCATCTGTTTCAATGCGAAACGATGCTGCTGCTGCTGCTGCT 381
 QY 961 GATCAGATCTGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1001
 Db 380 GATCAGATCTGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 340

RESULT 3

AA01440/c
 ID AA01440 standard; DNA; 3176 BP.

AC AA01440;

XX 28-APR-1999 (first entry)

XX Tn5seq1 transposon.

XX Tn5seq1 transposon; RNA transcription; gene hyperexpression;
 KW strong promoter; SP6 promoter; T7 promoter; ss.

XX Synthetic.

XX Key Location/Qualifiers

XX misc_feature 1500..1500

FT /*tag= a

FT /note= "specifically claimed region (claim 1)"

FT /*tag= b

FT /note= "O-end"

FT 3158..3176

FT /*tag= c

FT /note= "O-end"

XX US5869296-A.

XX

PD 09-FEB-1999.
 XX 14-JAN-1993; 93US-0004406.
 XX 05-OCT-1987; 87US-0105422.
 PR 12-APR-1990; 90US-0508382.
 PR 14-JAN-1993; 93US-0004406.
 XX (UNIW) UNIV WASHINGTON.
 PA Berg DE, Huang HV, Nag DK;
 PI WPI; 1999-152772/13.
 XX Obtaining hyperexpression of genes in Escherichia coli hosts - by
 PT insertion of transposon Tn5seq1 such that the strong SP6 and T7
 PT promoters are adjacent to the host genes
 XX Claim 1; Fig 4a-d; 19pp; English.
 XX This sequence represents the Tn5seq1 transposon used in the method of the
 CC invention. The method is for RNA transcription, and comprises the
 CC insertion of the Tn5seq1 transposon into an E. coli DNA molecule to
 CC obtain hyperexpression of genes adjacent to strong promoters SP6 or T7.
 CC The transposon is useful for stimulating the transcription of genes
 CC adjacent to the heterologous SP6 or T7 promoters in E. coli, for making
 CC RNA transcripts in vitro and the hyperexpression of specific
 CC transcription of genes (adjacent to the SP6 or T7 ends) in vivo. Tn5seq1
 CC offers a less laborious method of sequencing long DNA molecules than
 CC current methods such as base-specific chemical cleavage and enzymatic
 CC chain termination.
 XX Sequence 3176 BP; 661 A; 898 C; 893 G; 724 T; 0 other;
 SQ Query Match 100.0%; Score 1001; DB 20; Length 3176;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGCTGGAGGATCATCCAGCGGGTCCCGAAACAGTTCGGAAGCCCAACCTTTTCATAG 60
 Db 1340 CGCTGGAGGATCATCCAGCGGGTCCCGAAACAGTTCGGAAGCCCAACCTTTTCATAG 1281
 QY 61 AAGCGCGGTGGAATCGAAATCTCGTATGGAGGTGGGGTGGTGGTGGTGGTGGTGGT 120
 Db 1280 AAGCGCGGTGGAATCGAAATCTCGTATGGAGGTGGGGTGGTGGTGGTGGTGGTGGT 1221
 QY 121 TCGAACCCCGAGTCCCGCTCAGAAGAACTCGTCAAGAGCGGATAGAAGCGGATGCGCT 180
 Db 1220 TCGAACCCCGAGTCCCGCTCAGAAGAACTCGTCAAGAGCGGATAGAAGCGGATGCGCT 1161
 QY 181 GCGAATCGGAGCGGCGATACCGTAAAGCAGGAGCGGTGAGCCCATTCGCCGCCAA 240
 Db 1160 GCGAATCGGAGCGGCGATACCGTAAAGCAGGAGCGGTGAGCCCATTCGCCGCCAA 1101
 QY 241 GCTCTTCAGCAATATCAGCGGTAGCCACGCTATGCTGATAGCGGTGGCGCACACCCA 300
 Db 1100 GCTCTTCAGCAATATCAGCGGTAGCCACGCTATGCTGATAGCGGTGGCGCACACCCA 1041
 QY 301 GCCGGCCACAGTCGATGAATCCAGAAAGCGGCCATTTTCCACCATGATATTCGCAAGC 360
 Db 1040 GCCGGCCACAGTCGATGAATCCAGAAAGCGGCCATTTTCCACCATGATATTCGCAAGC 981
 QY 361 AGGATCCCATGTTGTCACGAGATCTTCGCGGTGGGCGATGCGGCTTGAGCGCTGG 420
 Db 980 AGGATCCCATGTTGTCACGAGATCTTCGCGGTGGGCGATGCGGCTTGAGCGCTGG 921
 QY 421 CGAACAGTTCGGTGGCGGAGCCCTGATGCTCTTCGTCAGATCATCTGATGACAA 480
 Db 920 CGAACAGTTCGGTGGCGGAGCCCTGATGCTCTTCGTCAGATCATCTGATGACAA 861
 QY 481 GACCGGTTCCATCCGAGTACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 Db 860 GACCGGTTCCATCCGAGTACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 801

```
QY 541 GGCAGGTAGCCGATCAAGCGTATGAGCGCGCGCATTCGATCAGCCATGATGGATCTT 600
Db 800 GGCAGGTAGCCGATCAAGCGTATGAGCGCGCGCATTCGATCAGCCATGATGGATCTT 741
QY 601 TCTCGGAGGAGCAAGGTGAGTACAGGAGATCTTCCCGCGGCACTTCGCCCAATAGCA 660
Db 740 TCTCGGAGGAGCAAGGTGAGTACAGGAGATCTTCCCGCGGCACTTCGCCCAATAGCA 681
QY 661 GGCAGTCCCTTCCGCTTCAGTGAACAGTGCAGACAGCTGCGAAGAAACGCCCGTCG 720
Db 680 GGCAGTCCCTTCCGCTTCAGTGAACAGTGCAGACAGCTGCGAAGAAACGCCCGTCG 621
QY 721 TGCCAGGACAGATAGCCGCTGCTCTGCTGAGTTCATTCAGGACCGGACAGGT 780
Db 620 TGCCAGGACAGATAGCCGCTGCTCTGCTGAGTTCATTCAGGACCGGACAGGT 561
QY 781 CGTCTTGAACAAAGAACCGGCGCCCTGCGCTGACAGCCGGAACACGCGGATCAG 840
Db 560 CGTCTTGAACAAAGAACCGGCGCCCTGCGCTGACAGCCGGAACACGCGGATCAG 501
QY 841 AGCAGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 500 AGCAGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 441
QY 901 GAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 440 GAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 381
QY 961 GATCAGATCTTATCCCTGCGCCATCAGATCTTGGCGGC 1001
Db 380 GATCAGATCTTATCCCTGCGCCATCAGATCTTGGCGGC 340
```

RESULT 4

AAQ27504/c
ID AAQ27504 standard; DNA; 3177 BP.

XX AAQ27504;

XX AC AAQ27504;

XX AC AAQ27504;

XX AC AAQ27504;

XX AC AAQ27504;

XX AC AAQ27504;

XX AC AAQ27504;

XX AC AAQ27504;

XX AC AAQ27504;

XX AC AAQ27504;

XX AC AAQ27504;

XX AC AAQ27504;

XX AC AAQ27504;

XX AC AAQ27504;

XX AC AAQ27504;

XX AC AAQ27504;

XX AC AAQ27504;

XX AC AAQ27504;

XX AC AAQ27504;

XX AC AAQ27504;

XX AC AAQ27504;

XX AC AAQ27504;

XX AC AAQ27504;

XX AC AAQ27504;

XX FN US5137829-A.
XX PD 11-AUG-1992.
XX PF 05-OCT-1987; 87US-0105422.
XX PR 05-OCT-1987; 87US-0105422.
XX PA (UNIW) UNIV WASHINGTON.
XX PI Berg DB, Huang HV, Nag DK;
XX DR WPI; 1992-292374/35.
XX PT New DNA transposon derived from transposon Tn5 - useful for
XX FT generation of mutants and rapid screening of DNA sequences
XX PS Claim 3; Fig 4; 19pp; English.
XX CC Transposon Tn5seq1 was constructed from the pBR322-derived plasmid
XX CC pBR322::Tn5. The novel transposon is contained in plasmid pBR322. It
XX CC can be used for sequencing bacterial genomic DNA without recombinant
XX CC DNA cloning and also for sequencing DNA cloned in phages or
XX CC multicopy plasmids. It may also be useful for RNA sequencing with
XX CC SP6 and T7 RNA polymerases, for making RNA transcripts in vitro and
XX CC for hyperexpression or specific transcription of adjacent genes in
XX CC vivo.
XX CC (Updated on 10-MAR-2003 to add missing OS field.)
XX SQ Sequence 3177 BP; 662 A; 898 C; 893 G; 724 T; 0 other;

Query Match

100.0%; Score 1001; DB 13; Length 3177;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGGAGGATCATCCAGCGCGGCTCCCGAAGAACGATTCGGAAGCCCAACCTTTTCATAG 60

Db 1340 CGCTGGAGGATCATCCAGCGCGGCTCCCGAAGAACGATTCGGAAGCCCAACCTTTTCATAG 1281

QY 61 AAGCGCGCGGTGGAATCGAAATCTGTGATGGCAGGTGGCGCTGCGTTCGTCGTCATT 120

Db 1280 AAGCGCGCGGTGGAATCGAAATCTGTGATGGCAGGTGGCGCTGCGTTCGTCGTCATT 1221

QY 121 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGCGATAGAGGCGATCGCT 180

Db 1220 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGCGATAGAGGCGATCGCT 1161

QY 181 GCGAATCGGAGCGGCGATACCGTAAAGCAAGAGGAGCGGTAGCCCAATTCGCCGCCAA 240

Db 1160 GCGAATCGGAGCGGCGATACCGTAAAGCAAGAGGAGCGGTAGCCCAATTCGCCGCCAA 1101

QY 241 GCTCTTCAGCAATATCAGCGGTAGCCAGCTATGCTCTGCTGATGCGGTCCGCGCACCCA 300

Db 1100 GCTCTTCAGCAATATCAGCGGTAGCCAGCTATGCTCTGCTGATGCGGTCCGCGCACCCA 1041

QY 301 GCCGCGCACAGTCCGATGAATCCAGAAAGCGGCCATTTTCCACCATGATATTCGCGCAAGC 360

Db 1040 GCCGCGCACAGTCCGATGAATCCAGAAAGCGGCCATTTTCCACCATGATATTCGCGCAAGC 981

QY 361 AGGCATCGCATGGGTACAGAGATCTCTCGCGTCCGCGCATTCGCGCTTTCGAGCTCG 420

Db 980 AGGCATCGCATGGGTACAGAGATCTCTCGCGTCCGCGCATTCGCGCTTTCGAGCTCG 921

QY 421 CGAAACAGTTTCGCTGGCGGAGCCCTGATGCTCTGCTGCTGATGCTCTGCTGATGCTGCTG 480

Db 920 CGAAACAGTTTCGCTGGCGGAGCCCTGATGCTCTGCTGCTGATGCTCTGCTGATGCTGCTG 861

QY 481 GACCGCTTCATCCAGTACGTCGCTCGATGCGATGCTTCGCTTCGCTTCGCTTCGCTTCG 540

Db 860 GACCGCTTCATCCAGTACGTCGCTCGATGCGATGCTTCGCTTCGCTTCGCTTCGCTTCG 801

QY 541 GGCAGGTAGCCGATCAAGCGTATGAGCGCGCGCATTCGATCAGCCATGATGGATCTT 600

Db 800 GGCAGGTAGCCGATCAAGGTATGACGCGCGCATTTGCATCAGCCATGATGATCTT 741
 Qy 601 TCTCGGACGAGCAAGGTAGATGACAGGAGATCTCTGCGCGCACTTCGCCCAATAGCA 660
 Db 740 TCTCGGACGAGCAAGGTAGATGACAGGAGATCTCTGCGCGCACTTCGCCCAATAGCA 681
 Qy 661 GCAGATCTCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT 720
 Db 680 GCAGATCTCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT 621
 Qy 721 TGCCACGACGATAGCCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 Db 620 TGCCACGACGATAGCCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 561
 Qy 781 CGTCTTGACAAAGAACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 Db 560 CGTCTTGACAAAGAACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 501
 Qy 841 AGCAGCGGATGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 Db 500 AGCAGCGGATGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 441
 Qy 901 GAGAACCTTGGTCAATCCATCTTGTTCATCATGCGAAACGATCTCATCTGCTCTT 960
 Db 440 GAGAACCTTGGTCAATCCATCTTGTTCATCATGCGAAACGATCTCATCTGCTCTT 381
 Qy 961 GATCAGATCTTGATCCCTTGGCGCATCAGATCTTGGCGGC 1001
 Db 380 GATCAGATCTTGATCCCTTGGCGCATCAGATCTTGGCGGC 340

RESULT 5

AA80959/c

ID AA80959 standard; DNA; 3740 BP.

XX AC AA80959;

XX 25-MAR-2003 (updated)

DT 12-OCT-1990 (first entry)

XX DE Plasmid pDML1 encoding neomycin resistance and lac repressor.

XX XX affinity peptide; metal-chelate affinity chromatography;

XW neomycin phosphotransferase (neo); lac repressor (lacI); pSM1,1; ss.

XX OS synthetic.

XX FT Key

FT CDS

FT Location/Qualifiers

FT 357..1151

FT /*tag= a

FT /product=neomycinphosphotransferase

FT 1432..2611

FT /*tag= b

FT /product=lac repressor

FT XX

FT XX

FT XX

FT XX

FT XX

FT XX

FT XX

FT XX

FT XX

FT XX

FT XX

FT XX

FT XX

FT XX

FT XX

FT XX

FT XX

FT XX

FT XX

FT XX

FT XX

FT XX

FT XX

FT XX

FT XX

FT XX

FT XX

XX

PS Disclosure; Page ?; pp; German.

XX

CC Plasmid encodes a fusion protein

CC consisting of 1 or 2 affinity peptides bonded directly or

CC indirectly to a biologically active protein (i.e. dihydrofolate

CC reductase). The presence of adjacent His residues in the affinity

CC peptides allows purification of the biologically active protein by

CC metal-chelate chromatography on nitrilotriacetic acid resins.

CC See also AAP0401-3 and AA80955-8 and AA80960-63.

CC (Updated on 25-MAR-2003 to correct PR field.)

XX

SQ Sequence 3740 BP; 813 A; 1035 C; 1056 G; 836 T; 0 other;

Query Match

Best Local Similarity 100.0%; Score 1001; DB 9; Length 3740;

Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

Db 1290 CGTGGAGGATCATCCAGCGCGCTCCGGAAACGATTCGGAAGCCCAACCTTTTCATAG 1231

Qy

Db 61 AAGCGCGCGTGAATCGAATCTCGTGATGGCAGGTTGGGCGTCTGCTTGGTCCGTCATT 120

Db

Qy 1230 AAGCGCGCGTGAATCGAATCTCGTGATGGCAGGTTGGGCGTCTGCTTGGTCCGTCATT 1171

Qy

Db 121 TCGAACCCAGAGTCCCGCTCAGAAGAACTCTGTCAGAGGCGGATAGAGGCGATGCGCT 180

Db

Qy 1170 TCGAACCCAGAGTCCCGCTCAGAAGAACTCTGTCAGAGGCGGATAGAGGCGATGCGCT 1111

Qy

Db 181 GCGAATCGGAGCGGATACCGTAAAGCAGAGGAGCGGTCAGCCCATTCGCCGCCAA 240

Db

Qy 1110 GCGAATCGGAGCGGATACCGTAAAGCAGAGGAGCGGTCAGCCCATTCGCCGCCAA 1051

Qy

Db 241 GCTCTTCAGCAATATCAOCSGTAGCAACGCTATGCTCTGATAGCGGTCCGCACACCA 300

Db

Qy 1050 GCTCTTCAGCAATATCAOCSGTAGCAACGCTATGCTCTGATAGCGGTCCGCACACCA 991

Qy

Db 301 GCGCGCCACAGTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATATTCGCAAGC 360

Db

Qy 990 GCGCGCCACAGTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATATTCGCAAGC 931

Qy

Db 361 AGCATTCGCATCGGTCAAGAGATCTTCGCGGTTCGCGCATTCGCGCTTGGAGCTGG 420

Db

Qy 930 AGCATTCGCATCGGTCAAGAGATCTTCGCGGTTCGCGCATTCGCGCTTGGAGCTGG 871

Qy

Db 421 CGAACAGTTCCGCTGGCGGCGCCCTGATGCTCTTCGTCAGATCATCTGATCGACAA 480

Db

Qy 870 CGAACAGTTCCGCTGGCGGCGCCCTGATGCTCTTCGTCAGATCATCTGATCGACAA 811

Qy

Db 481 GACCGGCTTCCATCCGAGTACGTCGCTCGATGCGATGCTTTCGCTTGGTGTGCAATG 540

Db

Qy 810 GACCGGCTTCCATCCGAGTACGTCGCTCGATGCGATGCTTTCGCTTGGTGTGCAATG 751

Qy

Db 541 GGCAGGTAGCGGATCAAGCGTATGAGCGCGCGCATTCGCATCAGCCATGATGATCTT 600

Db

Qy 750 GGCAGGTAGCGGATCAAGCGTATGAGCGCGCGCATTCGCATCAGCCATGATGATCTT 691

Qy

Db 601 TCTCGCAGGACAGGTGAGATGACAGAGATCTTCGCGCGCATTCGCCCAATAGCA 660

Db

Qy 690 TCTCGCAGGACAGGTGAGATGACAGAGATCTTCGCGCGCATTCGCCCAATAGCA 631

Qy

Db 661 GGCAGTCCCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT 720

Db

Qy 630 GGCAGTCCCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT 571

Qy

Db 721 TGCCACGACGATAGCCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780

Db

Qy 570 TGCCACGACGATAGCCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 511

Qy

Db 781 CGTCTTGACAAAGAACCGCGCGCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840

Db

Qy 510 CGTCTTGACAAAGAACCGCGCGCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 451

New fusion protein contg biologically active peptide -
 coupled to affinity peptide allowing purification by metal chelate
 affinity chromatography, useful eg in vaccines and as reagents

QY 841 AGCAGCCGATGCTGTTGTCGCCAGTCATAGCCGAATAGCCTCTCCACCCCAAGCGCGC 900
 Db 450 AGCAGCCGATGCTGTTGTCGCCAGTCATAGCCGAATAGCCTCTCCACCCCAAGCGCGC 391
 QY 901 GAGAACCTGCGTGCATCCTCTGTTCAATCATCGGAACGATCCTCATCTGTCTCTT 960
 Db 390 GAGAACCTGCGTGCATCCTCTGTTCAATCATCGGAACGATCCTCATCTGTCTCTT 331
 QY 961 GATCAGATCTTATCCCTGCGCATCAGATCCTTTGGCGGC 1001
 Db 330 GATCAGATCTTATCCCTGCGCATCAGATCCTTTGGCGGC 290

RESULT 6

AAN81154/c

ID AAN81154 standard; DNA; 3740 BP.

XX AC

XX AAN81154;

XX 25-MAR-2003 (updated)

DT 09-JAN-2003 (updated)

DT 12-NOV-1990 (first entry)

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

XX Plasmid pDM1.1.

QY 61 AAGCGCGGTCGATCGAAATCTCTGTATGCGAGGTCGCGCTCGCTTGGTCGTCATT 120
 Db 1230 AAGCGCGGTCGATCGAAATCTCTGTATGCGAGGTCGCGCTCGCTTGGTCGTCATT 1171
 QY 121 TCGAACCCAGAGTCGCGCTCAGAGAACTCGTCAGAGAGGCGATAGAGGCGATCGGCT 180
 Db 1170 TCGAACCCAGAGTCGCGCTCAGAGAACTCGTCAGAGAGGCGATAGAGGCGATCGGCT 1111
 QY 181 GCGAATCGGAGCGGCGATACCGTAAAGCAGAGGAGCGGTCAGCCCAATTGCGCCCAA 240
 Db 1110 GCGAATCGGAGCGGCGATACCGTAAAGCAGAGGAGCGGTCAGCCCAATTGCGCCCAA 1051
 QY 241 GCTCTTCAGCAATATCAGCGGTAGCCAAAGCTATGCTCTGTATGAGCGGTCCGACACCA 300
 Db 1050 GCTCTTCAGCAATATCAGCGGTAGCCAAAGCTATGCTCTGTATGAGCGGTCCGACACCA 991
 QY 301 GCGGCGCACAGTCGATGAATCCAGAAAAGCGGCAATTTCCACCATGATATTCGCAAGC 360
 Db 990 GCGGCGCACAGTCGATGAATCCAGAAAAGCGGCAATTTCCACCATGATATTCGCAAGC 931
 QY 361 AGSCATCGCATGGGTACAGAGATCTCTCGCGTCGCGCATGCGCGCTTGTAGCGCTGG 420
 Db 930 AGSCATCGCATGGGTACAGAGATCTCTCGCGTCGCGCATGCGCGCTTGTAGCGCTGG 871
 QY 421 CGAAACAGTTTCGCTGCGCGAGCCCTGATGCTCTCTGTCCAGATCATCTGATCCACAA 480
 Db 870 CGAAACAGTTTCGCTGCGCGAGCCCTGATGCTCTCTGTCCAGATCATCTGATCCACAA 811
 QY 481 GACCGCTTCATCCAGTACGTCGCTCGATCGATGTTTCGCTGCGTTCGTCGTAATG 540
 Db 810 GACCGCTTCATCCAGTACGTCGCTCGATCGATGTTTCGCTGCGTTCGTCGTAATG 751
 QY 541 GGCAGTACCGGATCAAGCGTATCGAGCGCGCGGATTCATCAGCCCATGATGATCTT 600
 Db 750 GGCAGTACCGGATCAAGCGTATCGAGCGCGCGGATTCATCAGCCCATGATGATCTT 691
 QY 601 TCTCGCAGGAGCAAGGTGAGATGACAGAGATCTCTGCCCCGCGCATTCGCCCAATAGCA 660
 Db 690 TCTCGCAGGAGCAAGGTGAGATGACAGAGATCTCTGCCCCGCGCATTCGCCCAATAGCA 631
 QY 661 GCGAGTCCCTTCGCGTTTCAGTGACAAAGTCAGACAGCTCGGACAGCTCGGACGCGCTG 720
 Db 630 GCGAGTCCCTTCGCGTTTCAGTGACAAAGTCAGACAGCTCGGACAGCTCGGACGCGCTG 571
 QY 721 TGGCCAGCCACGATAGCGCGCTGCTCTGCTCTGAGTTTCATTCAGGCGCACCGGACAGT 780
 Db 570 TGGCCAGCCACGATAGCGCGCTGCTCTGCTCTGAGTTTCATTCAGGCGCACCGGACAGT 511
 QY 781 CGGTCTTGACAAAAGAACCGGCGCGCTGCTGCTGAGCGCGGACACCGGCGGATCAG 840
 Db 510 CGGTCTTGACAAAAGAACCGGCGCGCTGCTGCTGAGCGCGGACACCGGCGGATCAG 451
 QY 841 AGCAGCGGATTCCTGTTGTCGCCAGTCATAGCCGAATAGCTCTCCACCCCAAGCGCGC 900
 Db 450 AGCAGCGGATTCCTGTTGTCGCCAGTCATAGCCGAATAGCTCTCCACCCCAAGCGCGC 391
 QY 901 GAGAACCTGCGTGCATCCTCTGTTCAATCATCGGAACGATCCTCATCTGTCTCTT 960
 Db 390 GAGAACCTGCGTGCATCCTCTGTTCAATCATCGGAACGATCCTCATCTGTCTCTT 331
 QY 961 GATCAGATCTTATCCCTGCGCATCAGATCCTTTGGCGGC 1001
 Db 330 GATCAGATCTTATCCCTGCGCATCAGATCCTTTGGCGGC 290

RESULT 7

AAN80483/c

ID AAN80483 standard; DNA; 3740 BP.

XX AC

XX AAN80483;

XX 25-MAR-2003 (updated)

DT 30-NOV-1990 (first entry)

QY 1 CGCTGAGGATCATCCAGCGCGTCCCGGAAAACGATTCCGAAAGCCCAACCTTTATAG 60
 Db 1290 CGCTGAGGATCATCCAGCGCGTCCCGGAAAACGATTCCGAAAGCCCAACCTTTATAG 1231

Query Match

Best Local Similarity 100.0%; Score 1001; DB 9; Length 3740;

Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Sequence 3740 BP; 813 A; 1036 C; 1055 G; 836 T; 0 other;

New immunogenic polypeptides - derived from Plasmodium falciparum 190 kD

surface antigen precursor, useful in vaccines, and encoding DNA sequences

Disclosure; Page ?; pp; German.

Plasmid pDM1.1 contains regions encoding neomycinophosphotransferase and

the lac repressor.

(Updated on 09-JAN-2003 to add missing OS field.)

(Updated on 25-MAR-2003 to correct PA field.)

(Updated on 25-MAR-2003 to correct PI field.)

XX pDM1.1 contg. the neo-gene of Tn5 and the lacI-gene.
 XX HTLV-III; env protein; gag protein; AIDS; immunogenic activity;
 KW vaccines; diagnosis; oligonucleotides; pDM1.1; neo-gene; lacI-gene; ss.
 XX Synthetic.
 XX Key Location/Qualifiers
 FT CDS 357..1148
 FT /*tag= a
 FT /product=neomycinphosphotransferase
 CDS 1532..2611
 FT /*tag= b
 FT /product=lac-repressor
 XX EP270114-A.
 XX 08-JUN-1988.
 XX 03-DEC-1987; 87EP-0117899.
 XX 05-DEC-1986; 86GB-0029116.
 XX (HOFF) HOFFMANN-LA ROCHE AG.
 XX Gentz R, Legrice S, Mous J, Stuber D;
 XX WPI; 1988-156343/23.
 XX New polypeptide with immunogenic activity of HTLV 3 env-gag protein
 FT - useful in vaccines and for AIDS diagnosis.
 XX Disclosure; Page ?; 49pp; German.
 XX Transposon Tn5: E. Beck et al., Gene 19, 327-336 (1982) for
 CC kanamycin resistance of E.coli.
 CC LacI-gene: P.J. Farabaugh, Nature 274, 765-769 (1978) contg. the
 CC mutation Iq (M.P. Calos, Nature 274, pp. 762-765 (1978) for the lac
 CC repressor.
 CC The plasmid also contains a region from plasmid pACYC184
 CC (A.C.Y. Chang and N.S. Cohen. J. Bacteriol. 134, 1141-1156 (1978) for
 CC stable replication.
 CC E.coli transfected with pDS5/RBSII.3A-5A (AAN80481), pDS8/RBSII
 CC (AAN80482) or derivatives, can'tg. a HTLV-gene, also contain plasmid
 CC pDM1.1. Expression is possible by addition of IPTG to the medium.
 CC gene is added.
 CC The DNA encodes peptides useful as active ingredients in vaccines
 CC and for detecting AIDS antibodies in serum etc.
 CC See also AAN80480-85.
 CC (Updated on 25-MAR-2003 to correct PD field.)
 CC (Updated on 25-MAR-2003 to correct PR field.)
 CC (Updated on 25-MAR-2003 to correct PI field.)
 CC (Updated on 25-MAR-2003 to correct DR field.)
 XX Sequence 3740 BP; 813 A; 1040 C; 1051 G; 836 T; 0 other;
 XX
 XX Query Match 100.0%; Score 1001; DB 9; Length 3740;
 XX Best Local Similarity 100.0%; Pred. NO. 0;
 XX Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1 CGCTGGAGGATCATCCAGCGGGTCCCGGAAACGATTCGAGCCCAACCTTTTCATAG 60
 Db 1290 CGCTGGAGGATCATCCAGCGGGTCCCGGAAACGATTCGAGCCCAACCTTTTCATAG 1231
 XX
 XX 61 AAGCGCGGTGGAATCGAAATCTCGTATGCGAGGTGGCGTGGCGTTCGTTGGTTCATT 120
 Db 1230 AAGCGCGGTGGAATCGAAATCTCGTATGCGAGGTGGCGTGGCGTTCGTTGGTTCATT 1171
 XX
 XX 121 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGCGATAGAGGCGATGCCCT 180
 Qy 1170 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGCGATAGAGGCGATGCCCT 1111
 Db

Qy 181 GCGAATCGGAGCGGCGATACCGTAAAGCAAGGAGCGGTGAGCCCATTCGCCGCGCAA 240
 Db 1110 GCGAATCGGAGCGGCGATACCGTAAAGCAAGGAGCGGTGAGCCCATTCGCCGCGCAA 1051
 Qy 241 GCTCTTCAGCAATATACGGGTAGCCAGCCTATGCTCTGATAGCGGTCCGACACCCA 300
 Db 1050 GCTCTTCAGCAATATACGGGTAGCCAGCCTATGCTCTGATAGCGGTCCGACACCCA 991
 Qy 301 GCCGGCCACAGTCGATGAATCCAGAAAGCGGCATTTTCCACCATGATATTCGGCAAGC 360
 Db 990 GCCGGCCACAGTCGATGAATCCAGAAAGCGGCATTTTCCACCATGATATTCGGCAAGC 931
 Qy 361 AGGCATCGCCATGGGTCAAGCAGATCTTCGGCTGCGGTGGGCATGCGGCTTGGAGCTGG 420
 Db 930 AGGCATCGCCATGGGTCAAGCAGATCTTCGGCTGCGGTGGGCATGCGGCTTGGAGCTGG 871
 Qy 421 CGAACAGTTCGGCTGGGCGGAGCCCTGATGCTCTTCGTCAGATCATCTGATCGACAA 480
 Db 870 CGAACAGTTCGGCTGGGCGGAGCCCTGATGCTCTTCGTCAGATCATCTGATCGACAA 811
 Qy 481 GACCGCTTCCATCCGAGTACGTCGCTCGTCGATGCGATGTTTTCGTTGGTGGTGAATG 540
 Db 810 GACCGCTTCCATCCGAGTACGTCGCTCGTCGATGCGATGTTTTCGTTGGTGGTGAATG 751
 Qy 541 GGCAGGTAGCCGATCAAGCGTATGAGCGCGGCTGATGCGCGGCGCATTCGATCGATGATCTT 600
 Db 750 GGCAGGTAGCCGATCAAGCGTATGAGCGCGGCTGATGCGCGGCGCATTCGATCGATGATCTT 691
 Qy 601 TCTCGGAGGAGCAAGGTGAGATGACAGGAGATCTTCCCGCGGCACTTCGCCCAATAGCA 660
 Db 690 TCTCGGAGGAGCAAGGTGAGATGACAGGAGATCTTCCCGCGGCACTTCGCCCAATAGCA 631
 Qy 661 GCCAGTCCCTTCCCGCTTTCAGTACACGTCGAGCAGCTGCGCAAGGAAACCCCGTCG 720
 Db 630 GCCAGTCCCTTCCCGCTTTCAGTACACGTCGAGCAGCTGCGCAAGGAAACCCCGTCG 571
 Qy 721 TGGCCAGCCAGATAGCCGCGCTCGCTTCCTGTCAGTTTCATTCAGGCGCACCGGACAGGT 780
 Db 570 TGGCCAGCCAGATAGCCGCGCTCGCTTCCTGTCAGTTTCATTCAGGCGCACCGGACAGGT 511
 Qy 781 CGCTCTTGACAAAGAAAGAACCGGGCGGCTTCGCTGACAGCGGCAACACGGCGGCATCAG 840
 Db 510 CGCTCTTGACAAAGAAAGAACCGGGCGGCTTCGCTGACAGCGGCAACACGGCGGCATCAG 451
 Qy 841 AGCAGCGGATGCTGTTGTGCGCCAGTCATAGCGGATAGCCTTCACCCAAAGCGGCGG 900
 Db 450 AGCAGCGGATGCTGTTGTGCGCCAGTCATAGCGGATAGCCTTCACCCAAAGCGGCGG 391
 Qy 901 GAGAACCTGCGTCAATTCATCTGTTCAATGCGAAACGATTCCTCATCTCTGCTCTTT 960
 Db 390 GAGAACCTGCGTCAATTCATCTGTTCAATGCGAAACGATTCCTCATCTCTGCTCTTT 331
 Qy 961 GATCAGATCTTGATCCCTGCGGCATCAGATCCCTTGGCGGC 1001
 Db 330 GATCAGATCTTGATCCCTGCGGCATCAGATCCCTTGGCGGC 290

RESULT 8
 AAQ06307/C
 ID AAQ06307 standard; DNA; 3740 BP.
 XX AAQ06307;
 AC AAQ06307;
 XX
 DT 25-MAR-2003 (updated)
 DT 29-JAN-1991 (first entry)
 XX
 DE Sequence of plasmid pDM1.1.
 XX
 KW IFN-gamma receptor; autoimmune disease; multiple sclerosis;
 KW hypersensitivity; ds.
 XX
 OS Homo sapiens.
 XX

PH Key Location/Qualifiers
 FT CDS 357..1146
 FT /*tag= a
 FT /label= Neomycin phosphotransferase.
 FT complement (1531..2609)
 FT /*tag= b
 FT /label= lac repressor.
 XX
 PN EP393502-A.
 XX
 XX 24-OCT-1990.
 XX
 XX 11-APR-1990; 90EP-0106992.
 XX
 XX 11-APR-1990; 90EP-0106992.
 XX
 PA (HOFF) HOFFMANN-LA ROCHE AG.
 XX
 XX Fountoulakis M, Garotta G, Stuber D;
 PI WPI; 1990-322042/43.
 XX
 DR Soluble interferon-gamma receptors - for treating auto-immune
 XX diseases, chronic inflammations, etc.
 XX
 XX Disclosure; Fig 16; 174pp; English.
 CC IFN-gamma is a therapeutically active agent in the treatment
 CC of autoimmune disease, allograft transplant rejections, multiple
 CC sclerosis, chronic inflammations and delayed hypersensitivity. It is
 CC also useful in identifying IFN-gamma agonists and antagonists.
 CC See also AAQ06301.
 CC (Updated on 25-MAR-2003 to correct PR field.)
 XX
 XX Sequence 3740 BP; 813 A; 1035 C; 1056 G; 836 T; 0 other;
 SQ

Query Match 100.0%; Score 1001; DB 11; Length 3740;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGGAGGATCATCCAGCGCGCTCCGGAAACAGATTCGAGCCCAACCTTTCATAG 60
 DB 1290 CGCTGGAGGATCATCCAGCGCGCTCCGGAAACAGATTCGAGCCCAACCTTTCATAG 1231
 QY 61 AAGCGCGCGTGGAAATCGAAATCTCGTGATGCGAGGTGGCGCTGCGTTCGTCATT 120
 DB 1230 AAGCGCGCGTGGAAATCGAAATCTCGTGATGCGAGGTGGCGCTGCGTTCGTCATT 1171
 QY 121 TCGRACCCAGATCCCGCTCAGAGAACTCGTCAAGAGCGGATGAGCGGATGCGCT 180
 DB 1170 TCGAACCCAGATCCCGCTCAGAGAACTCGTCAAGAGCGGATGAGCGGATGCGCT 1111
 QY 181 GCGAATCGGAGCGCGATACCGTAAAGCAGAGGAAACGCTCAGCCCATTCGCGCCAA 240
 DB 1110 GCGAATCGGAGCGCGATACCGTAAAGCAGAGGAAACGCTCAGCCCATTCGCGCCAA 1051
 QY 241 GCTCTTACGAAATATCAGGATGACCAACGCTATGCTGATAGCGGTCCGCCACACCCA 300
 DB 1050 GCTCTTACGAAATATCAGGATGACCAACGCTATGCTGATAGCGGTCCGCCACACCCA 991
 QY 301 GCCGGCCACAGTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATATTCGGCAAGC 360
 DB 990 GCCGGCCACAGTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATATTCGGCAAGC 931
 QY 361 AGGCATCGCCATGGGTCAAGCAGAGATCTTCGCGTCGGGATCGCGCTTCGAGCTGG 420
 DB 930 AGGCATCGCCATGGGTCAAGCAGAGATCTTCGCGTCGGGATCGCGCTTCGAGCTGG 871
 QY 421 CGAACAGTTCGGCTGGCGGCGGCCCTTGATGCTCTTCGTCAGATCATCTGATGACAA 480
 DB 870 CGAACAGTTCGGCTGGCGGCGGCCCTTGATGCTCTTCGTCAGATCATCTGATGACAA 811
 QY 481 GACCGGCTTCATCCGAGTAGTCTCGCTCGATGCGATGTTTCGCTTGGTGTGCAATG 540

DB 810 GACCGGCTTCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGTGCAATG 751
 QY 541 GCGAGGTAGCGGATCAAGCGTATGACGCGCGCGGATTCGATCAGCATGATGATGATCTT 600
 DB 750 GCGAGGTAGCGGATCAAGCGTATGACGCGCGCGGATTCGATCAGCATGATGATGATCTT 691
 QY 601 TCTCGCAGGAGCAAGGTGAGATGACAGGAGATCTTCGCGCGCGCACTTCGCGCCAAATAGCA 660
 DB 690 TCTCGCAGGAGCAAGGTGAGATGACAGGAGATCTTCGCGCGCGCACTTCGCGCCAAATAGCA 631
 QY 661 GCCAGTCCCTTCGCGCTTCAGTGACAACTGAGCAGCAGTGGCGGAGGAAAGCCCGTCG 720
 DB 630 GCCAGTCCCTTCGCGCTTCAGTGACAACTGAGCAGCAGTGGCGGAGGAAAGCCCGTCG 571
 QY 721 TGCCAGCCAGCATAGCCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 780
 DB 570 TGCCAGCCAGCATAGCCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 511
 QY 781 CGSTCTTGACAAAGAACCGCGCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 840
 DB 510 CGSTCTTGACAAAGAACCGCGCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 451
 QY 841 AGCAGCGGATTCGCTGCTGCGCGCTCATAGCCGAAATAGCTCTCCACCCCAAGCGCGCG 900
 DB 450 AGCAGCGGATTCGCTGCTGCGCGCTCATAGCCGAAATAGCTCTCCACCCCAAGCGCGCG 391
 QY 901 GAGAACCTCGTGCATTCATCTTGTTCATCATCGAAACGATCTCATCTGCTCTCTT 960
 DB 390 GAGAACCTCGTGCATTCATCTTGTTCATCATCGAAACGATCTCATCTGCTCTCTT 331
 QY 961 GATCAGATCTTCATCCCTCGCGCATCAGATCCTTGGCGGC 1001
 DB 330 GATCAGATCTTCATCCCTCGCGCATCAGATCCTTGGCGGC 290

RESULT 9
 AAQ12786/c
 ID AAQ12786 standard; DNA; 3740 BP.
 XX
 AC AAQ12786;
 XX
 DT 09-JAN-2003 (updated)
 DT 09-OCT-1991 (first entry)
 XX
 DE PDM1.1.
 XX
 KW Merozoite; antigen; vaccine; coccidiosis; RBS; ss.
 XX
 OS Unidentified.
 XX
 PN EP439056-A.
 XX
 PD 31-JUL-1991.
 XX
 PF 17-JAN-1991; 91EP-0100502.
 XX
 PR 26-JAN-1990; 90US-0470508.
 XX
 PA (HOFF) HOFFMANN-LA ROCHE AG.
 XX
 PI Binger MH;
 XX
 XX WPI; 1991-224487/31.
 DR
 XX Bimera merozoite surface antigen (23kD), DNA encoding it - and
 PT recombinant virus, useful in the prodn. of vaccine against
 PT coccidiosis in poultry
 XX
 PS Disclosure; Fig 11(a-c); 40pp; English.
 XX
 CC A 1.2 kb insert from lambda 5-7 was isolated, ends were repaired
 CC with Klenow polymerase and BamHI linkers were ligated to both ends.

CC The modified fragment was inserted into each of pDS56/RBSII,
 CC pDS56/RBSII-1, and pDS56/RBSII-2. Plasmids contg. the inserts
 CC in both possible orientations were transformed into E. coli
 CC strain M15 carrying the compatible pDM1.1. The E. coli strain
 CC M15 contg. plasmide pDS56/RBSII and pDM1.1 is described in
 CC EP-316695.
 CC See also AAQ12782-86.
 CC (Updated on 09-JAN-2003 to add missing OS field.)
 XX
 SQ Sequence 3740 BP; 809 A; 1039 C; 1056 G; 836 T; 0 other;

Query Match 100.0%; Score 1001; DB 12; Length 3740;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTGAGGATCATCCAGCGCGGTCCCGGAAACGATTCGGAAGCCCAACCTTTTCATAG 60
 DB 1290 CGTGTGAGGATCATCCAGCGCGGTCCCGGAAACGATTCGGAAGCCCAACCTTTTCATAG 1231

QY 61 AAGCGCGCGGTGGAATCGAATCTCGTGATGCGAGGTGGCGGTGCGTTCGCTCAT 120
 DB 1230 AAGCGCGCGGTGGAATCGAATCTCGTGATGCGAGGTGGCGGTGCGTTCGCTCAT 1171

QY 121 TCGAACCCAGAGTCCCGTTCGAGAACTCGTCAAGAGGCGATAGAGCGATGCGCT 180
 DB 1170 TCGAACCCAGAGTCCCGTTCGAGAACTCGTCAAGAGGCGATAGAGCGATGCGCT 1111

QY 181 GCGAATCGGAGCGCGATACCGTAAGACAGAGAGCGGTTCAGCCATTCGCGCGCAA 240
 DB 1110 GCGAATCGGAGCGCGATACCGTAAGACAGAGAGCGGTTCAGCCATTCGCGCGCAA 1051

QY 241 GCTCTTCAGCAATATCACGGGTAGCCAGCTATGCTGATAGCGGTCCGCGCACACCA 300
 DB 1050 GCTCTTCAGCAATATCACGGGTAGCCAGCTATGCTGATAGCGGTCCGCGCACACCA 991

QY 301 GCGCGCCAGTGCATGATTCAGAAAGAGCGGCATTTTCCACCATGATATTCGCGAAGC 360
 DB 990 GCGCGCCAGTGCATGATTCAGAAAGAGCGGCATTTTCCACCATGATATTCGCGAAGC 931

QY 361 AGGCATCGCATGGGTACAGCAGATCTCGCGTGGGCGATGCGCGCTTGAGCCCTGG 420
 DB 930 AGGCATCGCATGGGTACAGCAGATCTCGCGTGGGCGATGCGCGCTTGAGCCCTGG 871

QY 421 CGAACAGTTCGGTGGCGAGCCCTGATGCTTTCGTCAGATCATCTGATCGACAA 480
 DB 870 CGAACAGTTCGGTGGCGAGCCCTGATGCTTTCGTCAGATCATCTGATCGACAA 811

QY 481 GACCGGCTTCCATCCAGTACGTGCTCGATCGATGCTTTCGCTTGGTGGTGAATG 540
 DB 810 GACCGGCTTCCATCCAGTACGTGCTCGATCGATGCTTTCGCTTGGTGGTGAATG 751

QY 541 GCGAGTAGCGGATCAAGCGTATGCGCGCGCATTCATCAGCCATGATGATATCTT 600
 DB 750 GCGAGTAGCGGATCAAGCGTATGCGCGCGCATTCATCAGCCATGATGATATCTT 691

QY 601 TCTCGCAGGAGAGGTGATGATGACAGAGATCTGCGCGGACATTCGCGCCATAGCA 660
 DB 690 TCTCGCAGGAGAGGTGATGATGACAGAGATCTGCGCGGACATTCGCGCCATAGCA 631

QY 661 GCGAGTCCCTTCCCGTTCAGTGACAACTCGAGCAGATCGGCAAGGAGCGCCGCTCG 720
 DB 630 GCGAGTCCCTTCCCGTTCAGTGACAACTCGAGCAGATCGGCAAGGAGCGCCGCTCG 571

QY 721 TGCGCAGCAGATAGCGCGTCTGCTCTGAGATTCATTCAGGACCGGACACAGGT 780
 DB 570 TGCGCAGCAGATAGCGCGTCTGCTCTGAGATTCATTCAGGACCGGACACAGGT 511

QY 781 CGGTCTTGACAAAAGAACCGGCGCGCTCGCTGACAGCGGAGAACCGGCGGATCAG 840
 DB 510 CGGTCTTGACAAAAGAACCGGCGCGCTCGCTGACAGCGGAGAACCGGCGGATCAG 451

QY 841 AGCAGCGCATGTCTGTGTGCGCAAGTCTAGCCGAATAGCTCTCCACCCAAAGCGCGC 900

DB 450 AGCAGCGCATGTGTGTGTGCCCATGATAGCTCTCCACCAAGCGCGC 391
 QY 901 GAGAACCTCGTGCATTCATCTGTTCAATATCGGAACGATCTCATCCGTCTCTT 960
 DB 390 GAGAACCTCGTGCATTCATCTGTTCAATATCGGAACGATCTCATCCGTCTCTT 331
 QY 961 GATCAGATCTGATCCCTGCGGCATCAGATCTTGGCGC 1001
 DB 330 GATCAGATCTGATCCCTGCGGCATCAGATCTTGGCGC 290

RESULT 10

AAQ34610/C

ID AAQ34610 standard; DNA; 3740 BP.

XX AC AAQ34610;

XX AC

DT 25-MAR-2003 (updated)

DT 13-MAY-1993 (first entry)

XX XX

DE Plasmid pDM1.1.

XX XX

KW Plasmid; pDS56/RBSII; -1; -2; regulatable; promoter; operator;

KW N25OPN25OP29; ribosomal binding site; RBS; RBSII(-1);

KW RBSII(-2); Pq25; E. coli; phage; T5; lac; repressor; lacI; lacIQ;

KW pDM1.1; replication; beta-lactamase; terminator; lambda; CAT; T1;

KW chloramphenicol acetyltransferase; rrnS operon; reading frame;

KW neomycin phosphotransferase; Tn5; transposon; selection marker;

KW pACYC184; replication; transmission; ss.

XX XX

OS Synthetic.

XX XX

PN EP522482-A2.

XX XX

PD 13-JAN-1993.

XX XX

PF 06-JUL-1992; 92EP-0111407.

XX XX

PR 12-JUL-1991; 91US-0729099.

XX XX

PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX XX

FI Binger M, Pasamontes L;

XX XX

DR WPI; 1993-010336/02.

XX XX

PT Immunogenic specified polypeptide and fragments - capable of

PT inducing an immune response against eimeria parasites and free

PT from proteins produced by these parasites

XX XX

PS Disclosure; Fig 11; 72pp; English.

XX XX

CC This sequence represents the plasmid pDM1.1. This plasmid was used

CC in conjunction with the expression plasmids pDS56/RBSII, -1, and -2.

CC These plasmids contain the regulatable promoter/operator element

CC N25OPN25OP29 and the ribosomal binding sites (RBS) RBSII, RBSII(-1)

CC and RBSII(-2) respectively. These RBS are derived from the RBS of the

CC promoter Pq25 of the E. coli phage T5. These plasmids may be

CC maintained in E. coli only if the promoter/operator element is

CC repressed by the binding of a lac repressor to the operator. The lac

CC repressor is encoded by the lacI gene. N25OPN25OP29 can be repressed

CC efficiently only when a sufficient amount of repressor molecules are

CC present in the cells. Therefore the lacI allele, which contains a

CC promoter mutant responsible for increased expression of the repressor

CC gene was used. This lacI allele is present on the plasmid pDM1.1

CC pDM1.1 also contains the neomycin phosphotransferase gene from the

CC Tn5 transposon, which is used as a selection marker, and a region of

CC the plasmid pACYC184 which contains all the information required for

CC replication and stable transmission to daughter cells. The region of

CC these plasmids between bases 1-1120 contains the replication origin

CC and the gene for beta-lactamase. The beta-lactamase gene is modified

CC by elimination of the HincII and PstI cleavage sites. These plasmids

CC also contain cleavage sites for the enzymes SalI, PstI and HindIII,

CC the terminator of E. coli phage lambda, the promoter-free gene of
 CC chloramphenicol acetyltransferase (CAT) and the terminator T1 of the
 CC E. coli rmb operon. These plasmids differ by one nucleotide
 CC following the ATG start codon resulting in protein expression in all
 CC three potential reading frames.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 3740 BP; 813 A; 1035 C; 1056 G; 836 T; 0 other;

Query Match 100.0%; Score 1001; DB 14; Length 3740;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGAGGATCATCCAGCGGCTCCGGAAACGATTCGAGCCCAACCTTTCATAG 60
 DB 1290 CCTGAGGATCATCCAGCGGCTCCGGAAACGATTCGAGCCCAACCTTTCATAG 1231

QY 61 AAGCGCGGCTGGAATCGAATCTCGTGATGCGAGTTGGCGTGGCTTGGTTCGTCATT 120
 DB 1230 AAGCGCGGCTGGAATCGAATCTCGTGATGCGAGTTGGCGTGGCTTGGTTCGTCATT 1171

QY 121 TCGAACCCAGAGTCCCGCTCAGAAAGACTGTCAGAGCGGATAGAGCGATGCGCT 180
 DB 1170 TCGAACCCAGAGTCCCGCTCAGAAAGACTGTCAGAGCGGATAGAGCGATGCGCT 1111

QY 181 GCGAATCGGAGCGGCTATACGTAAGCAGAGGAGCGTTCAGCCATTTCGCCGCAA 240
 DB 1110 GCGAATCGGAGCGGCTATACGTAAGCAGAGGAGCGTTCAGCCATTTCGCCGCAA 1051

QY 241 GCTCTTACGATATACGGGTAGCCAAAGCTATGTCCTGATAGCGGTTCGCCCAACCCA 300
 DB 1050 GCTCTTACGATATACGGGTAGCCAAAGCTATGTCCTGATAGCGGTTCGCCCAACCCA 991

QY 301 GCGGCCACAGTCGATGAATCCAGAAAGCGCCATTTCCACCATGATATTCGCGAAGC 360
 DB 990 GCGGCCACAGTCGATGAATCCAGAAAGCGCCATTTCCACCATGATATTCGCGAAGC 931

QY 361 AGGCATCGGCTGCGTACGACGATCCTCGCGTCGGGATCGCGCTTGGACCTGG 420
 DB 930 AGGCATCGGCTGCGTACGACGATCCTCGCGTCGGGATCGCGCTTGGACCTGG 871

QY 421 CGAACAGTTCCGCTGGCGGAGCCCTGATCTTCTGTCAGATCATCTCTGATCGCAA 480
 DB 870 CGAACAGTTCCGCTGGCGGAGCCCTGATCTTCTGTCAGATCATCTCTGATCGCAA 811

QY 481 GACCGGCTTCCATCCGAGTACGTCGCTCGCTCGATGCGATGTTTCGCTTGGTTCGAATG 540
 DB 810 GACCGGCTTCCATCCGAGTACGTCGCTCGCTCGATGCGATGTTTCGCTTGGTTCGAATG 751

QY 541 GCGAGTACCGGATCAAGCTATGACGCGCGGATTCGATCAGCCATGATGATGATCTT 600
 DB 750 GCGAGTACCGGATCAAGCTATGACGCGCGGATTCGATCAGCCATGATGATGATCTT 691

QY 601 TCTCGGAGGAGCAGGTGAGATGACAGAGATCTCGCCCGGACCTTCGCCCAATAGCA 660
 DB 690 TCTCGGAGGAGCAGGTGAGATGACAGAGATCTCGCCCGGACCTTCGCCCAATAGCA 631

QY 661 GCGAGTCCCTCCGCTTCAAGTACGACAGTTCGACGACAGTTCGACGAGCAACGCCCTCG 720
 DB 630 GCGAGTCCCTCCGCTTCAAGTACGACAGTTCGACGACAGTTCGACGAGCAACGCCCTCG 571

QY 721 TGGCCAGCAGATAGCGGCTCGCTGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 DB 570 TGGCCAGCAGATAGCGGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 511

QY 781 CGGCTTTCACAAAAGAACCGGCGGCGCCCTCGCTGATGACGCGGAAACACGCGCGCATCAG 840
 DB 510 CGGCTTTCACAAAAGAACCGGCGGCGCCCTCGCTGATGACGCGGAAACACGCGCGCATCAG 451

QY 841 AGCAGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 DB 450 AGCAGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 391

QY 901 GAGAACCTCGTGCAATCCATCTTGTTCATCATCGAAACGATCCTCCTGTCTCTT 960
 DB 390 GAGAACCTCGTGCAATCCATCTTGTTCATCATCGAAACGATCCTCCTGTCTCTT 331

QY 961 GATCAGATCTTGATCCCTGCGCCATCAGATCTTGGCGGC 1001
 DB 330 GATCAGATCTTGATCCCTGCGCCATCAGATCTTGGCGGC 290

RESULT 11

AAV33721

ID AAV33721 standard; DNA; 3974 BP.

XX AC AAV33721;

XX DT 02-FEB-1999 (first entry)

XX Bacterial expression vector pHE4a.

XX Thymus receptor tyrosine kinase; TRTK; LERK; Eph; colon carcinoma;
 KW testicular carcinoma; pancreatic carcinoma; lung adenocarcinoma;
 KW breast carcinoma; hepatocellular carcinoma; cancer;
 KW cell proliferation; diagnosis; therapy; prognosis; human;
 KW vector; pHE4a; ds.

XX Chimeric - Escherichia coli.

OS Chimeric - Bacteriophage T5.

XX Key Location/Qualifiers

FH -35_signal 3889..3894

FT /tag= a

FT misc_signal 3895..3911

FT /tag= b

FT -10_signal 3912..3917

FT /function= operator 1

FT misc_signal 3925..3943

FT /tag= d

FT -10_signal 3957..3962

FT /function= operator 2

FT RBS 3957..3962

FT /tag= e

XX WO9844111-A1.

FN 08-OCT-1998.

XX 27-MAR-1998; 98WO-US06021.

XX 28-MAR-1997; 97US-0042856.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Soppet DR;

XX WPI; 1998-542702/46.

XX New isolated thymus receptor tyrosine kinase - is used to develop
 PT products for the diagnosis and treatment of cancers and other
 PT diseases involving aberrant cell proliferation

XX Disclosure; Page 76-79; 101pp; English.

XX This is the nucleotide sequence of bacterial expression vector
 CC pHE4a. This vector includes a neomycin phosphotransferase gene
 CC as a selectable marker, an Escherichia coli origin of replication
 CC (oriC from pUC129), a phage T5 promoter sequence, 2 lac operators
 CC (see also AAV33722), a Shine-Dalgarno sequence, the lactose operon
 CC repressor gene (lacIq) and a multiple cloning site linker region.
 CC Expression of operatively linked sequences located downstream from
 CC the lac operators may be induced by addition of a lac operon
 CC inducer such as IPTG. pHE4 vectors have been produced that include
 CC the coding sequence (see AAV33697) for novel human thymus receptor
 CC tyrosine kinase (TRTK, see AAV70525-26). These can be used for the

CC production of recombinant TRK polypeptides in bacterial host cells.
CC The invention provides TRK polynucleotides and polypeptides.
CC vectors, host cells and recombinant methods for producing these.
CC Also provided are diagnostic methods for detecting disease states
CC associated with the aberrant expression of TRK and therapeutic
CC methods for treating such disease states, especially cancer.
XX
SQ Sequence 3974 BP; 899 A; 1116 C; 1095 G; 864 T; 0 other;
Query Match 100.0%; Score 1001; DB 19; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CGGTGAGGATCATCCAGCGCGGTCCCGGAAACGATTCGAGCCCAACCTTTTCATAG 60
Db 2579 CGCTGGAGGATCATCCAGCGCGGTCCCGGAAACGATTCGAGCCCAACCTTTTCATAG 2638
Qy 61 AAGCGCGCGTGAATCGAAATCTCGTGATGCGAGGTGCGGTGCGTGTGCGTCAAT 120
Db 2639 AAGCGCGCGTGAATCGAAATCTCGTGATGCGAGGTGCGGTGCGTGTGCGTCAAT 2698
Qy 121 TCGAACCCAGAGTCCCGTTCAGAAAGTCTGTCAGAAAGGCGATAGAGGCGATGCGCT 180
Db 2699 TCGAACCCAGAGTCCCGTTCAGAAAGTCTGTCAGAAAGGCGATAGAGGCGATGCGCT 2758
Qy 181 GCGAATCGGAGCGGCGATACCGTAAAGCAGCAGGAGCGGTACGCCATTCGCCGCCAA 240
Db 2759 GCGAATCGGAGCGGCGATACCGTAAAGCAGCAGGAGCGGTACGCCATTCGCCGCCAA 2818
Qy 241 GCTCTTCAGCAATATCACGGGTAGCCAAAGCTATGTCCTGTATAGCGGTCCGCCACACCCA 300
Db 2819 GCTCTTCAGCAATATCACGGGTAGCCAAAGCTATGTCCTGTATAGCGGTCCGCCACACCCA 2878
Qy 301 GCGCGCCACAGTCGATCCAGAAAGCGGCCATTTCCACCATGATATTCGCGCAAGC 360
Db 2879 GCGCGCCACAGTCGATCCAGAAAGCGGCCATTTCCACCATGATATTCGCGCAAGC 2938
Qy 361 AGGCATCGCCATGGGTACGACGAGATCCTCGCGGTGCGGCATGCGCGCTTGACCTGG 420
Db 2939 AGGCATCGCCATGGGTACGACGAGATCCTCGCGGTGCGGCATGCGCGCTTGACCTGG 2998
Qy 421 CGAACAGTTCGGTGGCGGAGCCCTGATGCTCTTCGTCAGATCATCTGATCGACAA 480
Db 2999 CGAACAGTTCGGTGGCGGAGCCCTGATGCTCTTCGTCAGATCATCTGATCGACAA 3058
Qy 481 GACCGGCTTCATCCAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGTGTCGAATG 540
Db 3059 GACCGGCTTCATCCAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGTGTCGAATG 3118
Qy 541 GGCAGTAGCCGATCAAGCGTATGACGCGCGCATTCGATCAGCCATGATGATACCT 600
Db 3119 GGCAGTAGCCGATCAAGCGTATGACGCGCGCATTCGATCAGCCATGATGATACCT 3178
Qy 601 TCTCGCAGAGCAAGTGATGACAGAGATCCTGCGCGGCTTCGCGCATGACCA 660
Db 3179 TCTCGCAGAGCAAGTGATGACAGAGATCCTGCGCGGCTTCGCGCATGACCA 3238
Qy 661 GGCAGTCCCTTCCCGTTCAGTACAACTGACGACAGTCTGCGGAAAGACGCGCGTGG 720
Db 3239 GGCAGTCCCTTCCCGTTCAGTACAACTGACGACAGTCTGCGGAAAGACGCGCGTGG 3298
Qy 721 TGSCCAGCAGATAGCGCGGTGCTCTGCTGAGTTCATTCAGGACCGGACAGGT 780
Db 3299 TGSCCAGCAGATAGCGCGGTGCTCTGCTGAGTTCATTCAGGACCGGACAGGT 3358
Qy 781 CGGTCTTGACAAAGAACCGGCGCCCTCGCTGACAGCGGAAACAGCGGGGATCAG 840
Db 3359 CGGTCTTGACAAAGAACCGGCGCCCTCGCTGACAGCGGAAACAGCGGGGATCAG 3418
Qy 841 AGCAGCGGATGCTGTGTCCTGCTCATAGCCGAATAGCTCTCCACCCAAAGCGCGG 900
Db 3419 AGCAGCGGATGCTGTGTCCTGCTCATAGCCGAATAGCTCTCCACCCAAAGCGCGG 3478

Qy 901 GAGAACCTGCTGCAATCCATCTGTTCAATCATCGAAACGATCCTCATCTCTCTTT 960
Db 3479 GAGAACCTGCTGCAATCCATCTGTTCAATCATCGAAACGATCCTCATCTCTCTTT 3538
Qy 961 GATCAGATCTTGATCCCTGCGGCATCAGATCCTTGGCGGC 1001
Db 3539 GATCAGATCTTGATCCCTGCGGCATCAGATCCTTGGCGGC 3579
RESULT 12
AAV66372
ID AAV66372 standard; cDNA; 3974 BP.
XX
AC AAV66372;
XX
DT 18-JAN-1999 (first entry)
XX
DE Vector pHE4a.
XX
KW Cardiac and pancreatic protein; CAPP; muscle-derived growth factor;
KW human; pancreatitis; myocardial infarction; cardiomyopathy;
KW vector; pHE4a; ds.
XX
OS Chimeric - Escherichia coli.
OS Chimeric - bacteriophage T5.
OS Chimeric - synthetic.
XX
FN WO9844112-A1.
XX
PD 08-OCT-1998.
XX
PF 27-MAR-1998; 98WO-US06022.
XX
PR 28-MAR-1997; 97US-0042855.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Soppet DR;
XX
DR WPI; 1998-557111/47.
XX
XX
XX New isolated cardiac and pancreatic protein - used to develop
XX products for the diagnosis and treatment of e.g. pancreatitis or
XX abnormal hypertrophy of the heart
XX
XX Disclosure; Page 73-75; 113pp; English.
XX
XX This is the nucleotide sequence of plasmid pHE4a (ATCC 209645). It
XX includes a neomycin phosphotransferase gene as selectable marker,
XX an Escherichia coli origin of replication, a T5 phage promoter
XX sequence, 2 lac operator sequences, a Shine-Delgarno sequence,
XX the lactose operon repressor gene (lacIq) and a multiple cloning
XX site linker region. A nucleotide sequence (see AAV66372) coding for
XX novel human cardiac and pancreatic protein (CAPP, see AAV66373) of
XX be operatively linked to the promoter and operator (see AAV66373) of
XX pHE4a, and CAPP expressed in E. coli cells following IPTG
XX induction. CAPP polypeptides can modulate the differentiation and
XX proliferation of cells and tissue, both in vivo and ex vivo. The
XX products can be used in the diagnosis and treatment of pancreatitis
XX and conditions that cause abnormal hypertrophy of the heart, such
XX as hypertension, myocardial infarction, valve disease and
XX cardiomyopathy. The products can also be used in detection and
XX cell culturing.
SQ Sequence 3974 BP; 899 A; 1116 C; 1095 G; 864 T; 0 other;
Query Match 100.0%; Score 1001; DB 19; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CGGTGAGGATCATCCAGCGCGGTCCCGGAAACGATTCGAGCCCAACCTTTTCATAG 60
Db 2579 CGCTGGAGGATCATCCAGCGCGGTCCCGGAAACGATTCGAGCCCAACCTTTTCATAG 2638

QY 61 AAGGCGCGGTGGAATCGAAATCTCGTATGCGAGGTTGGCGTGGCTTGGTTCGCTCAT 120
DB 2639 AAGGCGCGGTGGAATCGAAATCTCGTATGCGAGGTTGGCGTGGCTTGGTTCGCTCAT 2698
QY 121 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGGCGATGCGCT 180
DB 2699 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGGCGATGCGCT 2758
QY 181 GCGAATCGGAGCGCGATACCGTAAAGCACAGAGAGCGGTACGCCATTCGCGCCAA 240
DB 2759 GCGAATCGGAGCGCGATACCGTAAAGCACAGAGAGCGGTACGCCATTCGCGCCAA 2818
QY 241 GCTCTTCAGCATATCAGCGGTAGCCACGCTATGCTGTATAGCGGTCCGCGACACCCA 300
DB 2819 GCTCTTCAGCATATCAGCGGTAGCCACGCTATGCTGTATAGCGGTCCGCGACACCCA 2878
QY 301 GCGGCGCACAGTCGATGATCCAGAAAGCGGCCATTTCCACCATGATATTCGGCAAGC 360
DB 2879 GCGGCGCACAGTCGATGATCCAGAAAGCGGCCATTTCCACCATGATATTCGGCAAGC 2938
QY 361 AGGCATCGCATGGGTACGACGAGATCTCGCTCGGCGATGCGCGCTTGAGCCTGG 420
DB 2939 AGGCATCGCATGGGTACGACGAGATCTCGCTCGGCGATGCGCGCTTGAGCCTGG 2998
QY 421 CGAACAGTTCGGCTGCGGCGAGCCCTGATGCTCTTCGTCAGATCATCTGATCGCAA 480
DB 2999 CGAACAGTTCGGCTGCGGCGAGCCCTGATGCTCTTCGTCAGATCATCTGATCGCAA 3058
QY 481 GACCGGCTTCATCCGAGTACGTGCTCGCTCGATGCTGTTTCGTTGGTGGTGAATG 540
DB 3059 GACCGGCTTCATCCGAGTACGTGCTCGCTCGATGCTGTTTCGTTGGTGGTGAATG 3118
QY 541 GCGAGTACCGGATCAGCGTATGCGCGCGCATGTCATCGACCATGATGATGATCT 600
DB 3119 GCGAGTACCGGATCAGCGTATGCGCGCGCATGTCATCGACCATGATGATGATCT 3178
QY 601 TCTCGGAGGAGCAAGGTGAGATGACAGGATCTCGCCCGGCACTTCGCCCAATAGCA 660
DB 3179 TCTCGGAGGAGCAAGGTGAGATGACAGGATCTCGCCCGGCACTTCGCCCAATAGCA 3238
QY 661 GCGAGTTCCTCCCGCTTCACTGACAGTACAGAGATCTCGCCCGGCACTTCGCCCAATAGCA 720
DB 3239 GCGAGTTCCTCCCGCTTCACTGACAGTACAGAGATCTCGCCCGGCACTTCGCCCAATAGCA 3298
QY 721 TGGCAGGACCATAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 3299 TGGCAGGACCATAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3358
QY 781' CGGTCTTGACAAAAAGAACCGGCGCCCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
DB 3359 CGGTCTTGACAAAAAGAACCGGCGCCCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3418
QY 841 AGCAGCGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
DB 3419 AGCAGCGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3478
QY 901 GAGAACCTGCGTGCATTCATTTGTTCAATCATSCGAAACGATTCCTCATCTGCTGCTT 960
DB 3479 GAGAACCTGCGTGCATTCATTCATTTGTTCAATCATSCGAAACGATTCCTCATCTGCTGCTT 3538
QY 961 GATCAGATTCGATCCCTCGGCGATCAGATCCTTGGCGGC 1001
DB 3539 GATCAGATTCGATCCCTCGGCGATCAGATCCTTGGCGGC 3579

RESULT 13

AAV56699

ID AAV56699 standard; cDNA; 3974 BP.

XX

AC AAV56699;

XX

DT 26-NOV-1998 (first entry)

XX Plasmid pHE4-5 cDNA.
DE
XX
KW ELL2; RNA polymerase II elongation factor; human; neoplastic disorder;
KW leukaemia; inhibitor; diagnosis; susceptibility; ds.
XX
OS Synthetic.
XX
FN WO9837194-A1.
XX
PD 27-AUG-1998.
XX
PF 19-FEB-1998; 98WO-US03177.
XX
PR 19-FEB-1997; 97US-0038447.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX
PI Conaway JW, Conaway RC, Duan DR, Shilatifard A;
XX
DR WPI; 1998-467557/40.
XX
XX ELL2 RNA polymerase II elongation factor - and antibodies; for the
PT prevention and treatment of neoplastic disorders, e.g. leukaemia
PT
XX
PS Disclosure; Page 61-63; 84pp; English.
XX
CC This sequence is plasmid pHE4-5 which is used in a method to isolate and
CC characterise a human RNA polymerase II elongation factor ELL2 which can
CC be used in the design of treatments for neoplastic disorders (especially
CC leukemias). Antagonists of the polypeptides or nucleic acids that
CC inhibit the expression of the protein, can be used to inhibit ELL2
CC activity in a patient. A method for diagnosing or determining a
CC susceptibility to neoplastic disorders comprises assaying ELL2 gene
CC expression level in mammalian cells or body fluid and comparing these
CC levels with a standard, such that an increase or decrease in levels
CC indicates an increased or decreased susceptibility. The gene expression
CC level is assayed using the antibody, or by detecting ELL2 mRNA levels.
CC The host cells of may be used to identify compounds which bind to an
CC ELL2 polypeptide.
XX
SQ Sequence 3974 BP; 899 A; 1116 C; 1095 G; 864 T; 0 other;
Query Match 100.0%; Score 1001; DB 19; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGCTGAGGATCATCCAGCCGCGCTCCGGAAGAAAGATTCGGAAGCCCAACCTTTCATAG 60
DB 2579 CGCTGAGGATCATCCAGCCGCGCTCCGGAAGAAAGATTCGGAAGCCCAACCTTTCATAG 2638
QY 61 AAGGCGCGGTGGAATCGAAATCTCGTATGCGAGGTTGGCGTGGCTTGGTTCGCTCAT 120
DB 2639 AAGGCGCGGTGGAATCGAAATCTCGTATGCGAGGTTGGCGTGGCTTGGTTCGCTCAT 2698
QY 121 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGGCGATGCGCT 180
DB 2699 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGGCGATGCGCT 2758
QY 181 GCGAATCGGAGCGCGATACCGTAAAGCACAGAGAGCGGTACGCCATTCGCGCCAA 240
DB 2759 GCGAATCGGAGCGCGATACCGTAAAGCACAGAGAGCGGTACGCCATTCGCGCCAA 2818
QY 241 GCTCTTCAGCATATCAGCGGTAGCCACGCTATGCTGTATAGCGGTCCGCGACACCCA 300
DB 2819 GCTCTTCAGCATATCAGCGGTAGCCACGCTATGCTGTATAGCGGTCCGCGACACCCA 2878
QY 301 GCGGCGCACAGTCGATGATCCAGAAAGCGGCCATTTCCACCATGATATTCGGCAAGC 360
DB 2879 GCGGCGCACAGTCGATGATCCAGAAAGCGGCCATTTCCACCATGATATTCGGCAAGC 2938
QY 361 AGGCATCGCATGGGTACGACGAGATCTCGCTCGGCGATGCGCGCTTGAGCCTGG 420

Db 2939 AGGCATCGCCATGGGTCCAGCAGAGATCTCGCCGCTCGGCATCGCCGCTTGAGCCCTGG 2998
QY 421 CGAACAGTTTCGGTGGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 480
Db 2999 CGAACAGTTTCGGTGGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 3058
QY 481 GACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGTTGGTGGTGAATG 540
Db 3059 GACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGTTGGTGGTGAATG 3118
QY 541 GGCAGTACCGGATCAAGGATGATGACGCGCGGCAATTCATCAGCCATGATGATGAT 600
Db 3119 GGCAGTACCGGATCAAGGATGATGACGCGCGGCAATTCATCAGCCATGATGATGAT 3178
QY 601 TCTCGCAGGAGCAAGGTGAGATGACAGGAGATCTTCCGCGGCACTTCGCCCAATAGCA 660
Db 3179 TCTCGCAGGAGCAAGGTGAGATGACAGGAGATCTTCCGCGGCACTTCGCCCAATAGCA 3238
QY 661 GGCAGTACCGGATCAAGGATGATGACGCGCGGCAATTCATCAGCCATGATGATGAT 720
Db 3239 GGCAGTACCGGATCAAGGATGATGACGCGCGGCAATTCATCAGCCATGATGATGAT 3298
QY 721 TGGCCAGGACATAGCCGCTGCTCGCTCGATGATTCATTCAGGCGACCGGACAGGT 780
Db 3299 TGGCCAGGACATAGCCGCTGCTCGCTCGATGATTCATTCAGGCGACCGGACAGGT 3358
QY 781 CGGCTTTGACAAAGAACCGGCGCCCTGCGCTGACGACGACGCGGACGCGGATCAG 840
Db 3359 CGGCTTTGACAAAGAACCGGCGCCCTGCGCTGACGACGACGCGGACGCGGATCAG 3418
QY 841 AGCAGCGGATGTCTGTGTCGCTGCTCGCTCGATGATTCATTCAGGCGACCGGACAGGT 900
Db 3419 AGCAGCGGATGTCTGTGTCGCTGCTCGCTCGATGATTCATTCAGGCGACCGGACAGGT 3478
QY 901 GAGAACCTCGGTCAATCCATCTTGTTCATATGCGAAACGATCTTCCACCCAGCGGCG 960
Db 3479 GAGAACCTCGGTCAATCCATCTTGTTCATATGCGAAACGATCTTCCACCCAGCGGCG 3538
QY 961 GATCAGATCTTGATCCCTCGGCGATCAGATCTTGGCGG 1001
Db 3539 GATCAGATCTTGATCCCTCGGCGATCAGATCTTGGCGG 3579

RESULT 14

AAV57971
ID AAV57971 standard; cDNA; 3974 BP.

XX AAV57971;

AC AAV57971;

XX 19-NOV-1998 (first entry)

XX Vector pHE4-5 containing human MOGP.

XX Human; myelin oligodendrocyte glycoprotein-like protein; MOGP;
XX immunoglobulin superfamily; Ig; cancer; inflammation; diagnosis;
XX detection; multiple sclerosis; demyelinating disease; myelin; ds.

XX Homo sapiens.

XX Synthetic.

XX WO9833912-A1.

XX 06-AUG-1998.

XX 30-JAN-1998; 98WO-US01593.

XX 30-JAN-1997; 97US-0035445.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Chopra A, Gentz RL, Olsen HS, Ruben SM;

XX

DR WP1; 1998-437466/37.
XX Human myelin oligodendrocyte glycoprotein-like protein - used to
PT develop products for the diagnosis, prognosis and treatment of
PT cancers, inflammation or demyelinating disease such as multiple
PT sclerosis
XX
XX Disclosure; Page 63-65; 100pp; English.
XX
XX The present sequence represents pHE405 vector which contains human
CC myelin oligodendrocyte glycoprotein-like protein (MOGP). The presence
CC of MOGP on the surface of oligodendrocytes is associated with
CC oligodendrocyte maturation and with the production of myelin. Products
CC from the present invention can be used for the diagnosis, prognosis and
CC treatment of e.g. cancers, inflammation, or demyelinating disease such
CC as multiple sclerosis.
XX
XX Sequence 3974 BP; 899 A; 1116 C; 1095 G; 864 T; 0 other;

Query Match 100.0%; Score 1001; DB 19; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGAGGATCATCCAGCGCGCTCCGGAACCGATTCCGAAGCCCAACCTTTTCATAG 60
Db 2579 CGCTGAGGATCATCCAGCGCGCTCCGGAACCGATTCCGAAGCCCAACCTTTTCATAG 2638
QY 61 AAGCGCGGTGAATCGAAATCTCTGATGAGCAGGTTGGCGTCTGCTTGGTGGTCAAT 120
Db 2639 AAGCGCGGTGAATCGAAATCTCTGATGAGCAGGTTGGCGTCTGCTTGGTGGTCAAT 2698
QY 121 TCGAACCCCAAGTCCCGCTCAGAACTCTGTCAGAAAGCGATAGAGGCGATCGCT 180
Db 2699 TCGAACCCCAAGTCCCGCTCAGAACTCTGTCAGAAAGCGATAGAGGCGATCGCT 2758
QY 181 GCGAATCGGAGCGCGATACCGTAAAGCAGGAGGCGTCCAGCCCAATTCGCCGCCAA 240
Db 2759 GCGAATCGGAGCGCGATACCGTAAAGCAGGAGGCGTCCAGCCCAATTCGCCGCCAA 2818
QY 241 GCTCTTCAGCAATATCAGCGGTAGCAACGCTATGCTCTGATAGCGTCCGCCACACCA 300
Db 2819 GCTCTTCAGCAATATCAGCGGTAGCAACGCTATGCTCTGATAGCGTCCGCCACACCA 2878
QY 301 GCGCGCCACAGTCAATGAAATCCAGAAAGCGGCAATTTCCACCATGATATTCGGAAGC 360
Db 2879 GCGCGCCACAGTCAATGAAATCCAGAAAGCGGCAATTTCCACCATGATATTCGGAAGC 2938
QY 361 AGGCATCGCCATGGTTCAGCAGAGATCTTCGCCGTGGGCGATCGCGGCTTGAGCTGG 420
Db 2939 AGGCATCGCCATGGTTCAGCAGAGATCTTCGCCGTGGGCGATCGCGGCTTGAGCTGG 2998
QY 421 CGAACAGTTTCGGCTGGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 480
Db 2999 CGAACAGTTTCGGCTGGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 3058
QY 481 GACCGGTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGTTGGTGGTGAATG 540
Db 3059 GACCGGTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGTTGGTGGTGAATG 3118
QY 541 GGCAGGTAGCGGATCAAGCGTATGACGCGCGCGCATTCGATCAGCAGATGATGATGAT 600
Db 3119 GGCAGGTAGCGGATCAAGCGTATGACGCGCGCGCATTCGATCAGCAGATGATGATGAT 3178
QY 601 TCTCGCAGGAGCAAGGTGAGATGACAGGAGATCTTCCGCGGCACTTCGCCCAATAGCA 660
Db 3179 TCTCGCAGGAGCAAGGTGAGATGACAGGAGATCTTCCGCGGCACTTCGCCCAATAGCA 3238
QY 661 GGCAGTCCCTTCCGCTTCCAGTACAGCTGACAGCAGTCCGCGGAGGAGGAGGAGGAGG 720
Db 3239 GGCAGTCCCTTCCGCTTCCAGTACAGCTGACAGCAGTCCGCGGAGGAGGAGGAGGAGG 3298
QY 721 TGSCAGCCACGATAGCCGCTGCTCTGTCAGTTTCAATTCAGGCGACCGGACAGGT 780

Qy 841 AGCAGCCGATTGTCTGTGTGCCCCAGTCATAGCCGAATAGCCTCTCCACCCCAAGCGGCG 900
|||
Db 3419 AGCAGCCGATTGTCTGTGTGCCCCAGTCATAGCCGAATAGCCTCTCCACCCCAAGCGGCG 3478
|||
Qy 901 GAGAACCTGCGTGCAATCCATCTTGTTCATCATGCGAAACGATCCTCATCTGTCTCTT 960
|||
Db 3479 GAGAACCTGCGTGCAATCCATCTTGTTCATCATGCGAAACGATCCTCATCTGTCTCTT 3538
|||
Qy 961 GATCAGATCTTGATCCCGCTGCGGCATCAGATCCCTTGCGGCG 1001
|||
Db 3539 GATCAGATCTTGATCCCGCTGCGGCATCAGATCCCTTGCGGCG 3579
|||

Search completed: February 15, 2004, 22:40:31
Job time : 329.085 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 10:22:55 ; Search time 73.6147 Seconds
(without alignments)
6001.851 Million cell updates/sec

Title: US-09-921-143-36_COPY_4000_5000
Perfect score: 1001
Sequence: 1 CGCTGGAGGATCATCCAGCGGGTCCCGAAACGATTCGGAACCCACCTTTTCATAG 60

Scoring table: OLIGO_NUC
Gapop 60.0 , Capext 60.0

Searched: 569978 seqs, 220691566 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2.6/prodata/2/ina/5A-COMB.seq:*

2: /cgn2.6/prodata/2/ina/5B-COMB.seq:*

3: /cgn2.6/prodata/2/ina/6A-COMB.seq:*

4: /cgn2.6/prodata/2/ina/6B-COMB.seq:*

5: /cgn2.6/prodata/2/ina/PCUTUS-COMB.seq:*

6: /cgn2.6/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1001	100.0	2220	2	US-08-864-224-1
2	1001	100.0	3176	6	Sequence 1, Appli Patent No. 5212080-1
3	1001	100.0	3974	3	US-09-026-343-33
4	1001	100.0	3974	3	US-09-042-105-16
5	1001	100.0	3974	3	US-09-044-856A-7
6	1001	100.0	3974	3	US-09-023-082A-147
7	1001	100.0	3974	3	US-09-044-855A-7
8	1001	100.0	3974	3	US-09-078-670-4
9	1001	100.0	3974	4	US-09-026-408-14
10	1001	100.0	3974	4	US-09-362-871-33
11	1001	100.0	3974	4	US-09-627-154-4
12	1001	100.0	3974	4	US-09-027-287-50
13	1001	100.0	3974	4	US-09-437-602-4
14	1001	100.0	3974	4	US-09-252-658B-50
15	1001	100.0	3984	3	US-09-044-798A-10
16	1001	100.0	3984	4	US-09-725-460A-10
17	1001	100.0	4208	3	US-09-044-796A-9
18	1001	100.0	4208	3	US-09-725-460A-9
19	1001	100.0	4208	3	US-09-689-693-37
20	1001	100.0	4256	3	US-08-995-156A-21
21	1001	100.0	4256	4	US-09-419-281-21
22	1001	100.0	4277	3	US-09-044-796A-11
23	1001	100.0	4277	4	US-09-725-460A-11
24	1001	100.0	4328	3	US-09-132-808-1
25	1001	100.0	4328	3	US-08-910-647-2
26	1001	100.0	4328	4	US-09-620-925-2
27	1001	100.0	4328	4	US-09-620-260-1

28 1001 100.0 4328 4 US-09-620-259-1 Sequence 1, Appli

29 1001 100.0 4800 4 US-09-554-929-1 Sequence 1, Appli

30 1001 100.0 4818 3 US-08-910-647-4 Sequence 4, Appli

31 1001 100.0 4818 4 US-09-620-925-4 Sequence 4, Appli

32 1001 100.0 5107 3 US-08-910-647-3 Sequence 3, Appli

33 1001 100.0 5107 4 US-09-620-925-3 Sequence 3, Appli

34 1001 100.0 6561 4 US-09-380-190A-30 Sequence 30, Appli

35 1001 100.0 8068 4 US-09-301-593-27 Sequence 27, Appli

36 1001 100.0 8068 4 US-09-301-593-35 Sequence 35, Appli

37 1001 100.0 8349 4 US-09-186-002-16 Sequence 16, Appli

38 979 97.8 4958 4 US-09-380-190A-20 Sequence 20, Appli

39 979 97.8 5754 4 US-09-380-190A-18 Sequence 18, Appli

40 979 97.8 5754 4 US-09-380-190A-19 Sequence 19, Appli

41 979 97.8 6225 4 US-09-380-190A-1 Sequence 1, Appli

42 979 97.8 6321 4 US-09-380-190A-17 Sequence 17, Appli

43 979 97.8 6359 4 US-09-380-190A-15 Sequence 15, Appli

44 979 97.8 6795 4 US-09-380-190A-22 Sequence 22, Appli

45 979 97.8 6891 4 US-09-380-190A-16 Sequence 16, Appli

ALIGNMENTS

RESULT 1
US-08-864-224-1
; Sequence 1, Application US/08864224
; Patent No. 5851808
; GENERAL INFORMATION:
; APPLICANT: Ellledge, Stephen J.
; APPLICANT: Liu, Qinghua
; TITLE OF INVENTION: Rapid Subcloning Using Site-Specific
; TITLE OF INVENTION: Recombination
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/864,224
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: BCM-02681
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2220 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA"
US-08-864-224-1

Query Match 100.0%; Score 1001; DB 2; Length 2220;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CGCTGGAGGATCATCCAGCGGGTCCCGAAACGATTCGGAACCCACCTTTTCATAG 60
Db 932 CGCTGGAGGATCATCCAGCGGGTCCCGAAACGATTCGGAACCCACCTTTTCATAG 991

Qy	61	AAGCGCGCGTGC	AAATCGCAATCTCG	ATGATCGCAGGTTGG	CGCGCTTGGT	CGGTCAATT	120
Db	992	AAGCGCGCGTGC	AAATCGCAATCTCG	ATGATCGCAGGTTGG	CGCGCTTGGT	CGGTCAATT	1051
Qy	121	TCGAACCCCA	GAGTCCGCTC	AGAAGAACTCGT	CAAGAAGCGGAT	GAAGCGGATGCGCT	1111
Db	1052	TCGAACCCCA	GAGTCCGCTC	AGAAGAACTCGT	CAAGAAGCGGAT	GAAGCGGATGCGCT	1111
Qy	181	GCGAATCGG	AGCGCGGATAC	CGGTAAAGCAG	CAGGAAAGCGGT	CAGCCATTGCCGCGCAA	240
Db	1112	GCGAATCGG	AGCGCGGATAC	CGGTAAAGCAG	CAGGAAAGCGGT	CAGCCATTGCCGCGCAA	1171
Qy	241	GCTCTTCAG	CAATATCACGGGT	AGCCAAACGTAT	GTCTGATAGCGGT	TCGCGCACACCCA	300
Db	1172	GCTCTTCAG	CAATATCACGGGT	AGCCAAACGTAT	GTCTGATAGCGGT	TCGCGCACACCCA	1231
Qy	301	GC CGG CCA	CGATGATGAA	TCGAAAGCGGCCA	TATTTCA	CGATGATATTCGGCAAGC	360
Db	1232	GC CGG CCA	CGATGATGAA	TCGAAAGCGGCCA	TATTTCA	CGATGATATTCGGCAAGC	1291
Qy	361	AGGCATCGC	ATGSGGTAC	CAGCAGAGAT	TCCTCGCGT	CGGCATCGCGCTG	420
Db	1292	AGGCATCGC	ATGSGGTAC	CAGCAGAGAT	TCCTCGCGT	CGGCATCGCGCTG	1351
Qy	421	CGAACAGTT	CGGTCGCG	CGAGCCCTTG	ATGCTCTTCGTC	CCAGATCATCTGATCGACAA	480
Db	1352	CGAACAGTT	CGGTCGCG	CGAGCCCTTG	ATGCTCTTCGTC	CCAGATCATCTGATCGACAA	1411
Qy	481	GACCGGCTT	CCATCCGAGT	AGTGCTCGCT	CGATGCGAT	TTTCGTTGGTGGTGGAAATG	540
Db	1412	GACCGGCTT	CCATCCGAGT	AGTGCTCGCT	CGATGCGAT	TTTCGTTGGTGGTGGAAATG	1471
Qy	541	GGCAGGTAG	CGCGATCA	AGCGTATG	CAGCGCGCGCAT	TGTCATCAGCCATGATGGATACATT	600
Db	1472	GGCAGGTAG	CGCGATCA	AGCGTATG	CAGCGCGCGCAT	TGTCATCAGCCATGATGGATACATT	1531
Qy	601	TC TC GCG	CAGGACAA	GGTGGATG	ACAGGAGAT	TCCTGCCCGGCATCTGCCCCAAATAGCA	660
Db	1532	TC TC GCG	CAGGACAA	GGTGGATG	ACAGGAGAT	TCCTGCCCGGCATCTGCCCCAAATAGCA	1591
Qy	661	GCCAGTCCCT	TCCGCTTC	AGTCAACAGT	TCGAGCAG	CAGTGGCGAAGAACGCCCGTGG	720
Db	1592	GCCAGTCCCT	TCCGCTTC	AGTCAACAGT	TCGAGCAG	CAGTGGCGAAGAACGCCCGTGG	1651
Qy	721	TGCGCCAGC	CAAGTATAG	CGCGCTCGCT	CGCTTCGAGTT	CATTTCAGGGCACCGGCAGAGT	780
Db	1652	TGCGCCAGC	CAAGTATAG	CGCGCTCGCT	CGCTTCGAGTT	CATTTCAGGGCACCGGCAGAGT	1711
Qy	781	CGGTCTTTG	CAAAAAGAA	CCGGCGCCCT	TCGCGTCA	CAGCGGAAACAGCGGCGCATCAG	840
Db	1712	CGGTCTTTG	CAAAAAGAA	CCGGCGCCCT	TCGCGTCA	CAGCGGAAACAGCGGCGCATCAG	1771
Qy	841	AGCAGCGGA	TTGTCGTGTG	CGCCAGTCA	TAGCCGAATAG	CCCTCTCCACCCAGCGGCGG	900
Db	1772	AGCAGCGGA	TTGTCGTGTG	CGCCAGTCA	TAGCCGAATAG	CCCTCTCCACCCAGCGGCGG	1831
Qy	901	GAGAACCTG	CGGTGCAAT	CTTGTGTTCA	ATCATGCGAAAC	AGATCTCTCATCTCTGTCTCTT	960
Db	1832	GAGAACCTG	CGGTGCAAT	CTTGTGTTCA	ATCATGCGAAAC	AGATCTCTCATCTCTGTCTCTT	1891
Qy	961	GATCAGAT	CTTGAATCC	CTTGGCCAT	CAGATC	CTTGGCGG	1001
Db	1892	GATCAGAT	CTTGAATCC	CTTGGCCAT	CAGATC	CTTGGCGG	1932

RESULT 2

5212080-1/c

;Patent No. 5212080

; APPLICANT: NAG, DILIP K.; HUANG, HENRY V.; BERG, DOUGLAS E.

; TITLE OF INVENTION: METHOD OF DNA SEQUENCING USING DNA

;TRANSPONSON TN5SEQL

; NUMBER OF SEQUENCES: 6

QY 901 GAGAACCTTCGCTGCAATCCATCTTCTTCAATCATGCGAAACGATCCCTCATCTCTCTT 960
DB 440 GAGAACCTTCGCTGCAATCCATCTTCTTCAATCATGCGAAACGATCCCTCATCTCTT 381
QY 961 GATCAGATCTTGCATCCCTGCGCCATCAGATCCCTTGGCGGC 1001
DB 380 GATCAGATCTTGCATCCCTGCGCCATCAGATCCCTTGGCGGC 340

RESULT 3

US-09-026-343-33
; Sequence 33, Application US/09026343
; Patent No. 6008018
; GENERAL INFORMATION:
; APPLICANT: DUAN, D. ROXANNE
; APPLICANT: SHILATIFARD, ALI
; APPLICANT: CONAWAY, JOAN W.
; APPLICANT: CONAWAY, RONALD C.
; TITLE OF INVENTION: ELL2, A New Member of an ELL Family of
; RNA Polymerase II Elongation Factors
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/038,447
; FILING DATE: 19-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: GOLDSTEIN, JORGE A.
; REGISTRATION NUMBER: 29,021
; REFERENCE/DOCKET NUMBER: 1488.0880001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 372-2540
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3974 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cdna
US-09-026-343-33

Query Match 100.0%; Score 1001; DB 3; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGCTGGAGGATCATCCAGCGCGCTCCGGAACCGATCCGAAAGCCCAACCTTTTCATAG 60
DB 2579 CGCTGGAGGATCATCCAGCGCGCTCCGGAACCGATCCGAAAGCCCAACCTTTTCATAG 2638
QY 61 AAGCGCGGTGGATCGAATCTCGTATGGCAGAGTTGGCGCTCGCTTGGTCGTCATT 120
DB 2639 AAGCGCGGTGGATCGAATCTCGTATGGCAGAGTTGGCGCTCGCTTGGTCGTCATT 2698
QY 121 TCGAACCCAGAGTCCCGCTCAGAGAACCTCGTCAAGAGCGGATAGAGGCGATCGCT 180
DB 2699 TCGAACCCAGAGTCCCGCTCAGAGAACCTCGTCAAGAGCGGATAGAGGCGATCGCT 2758

QY 181 GCGAATCGGAGCGGCGATACCGTAAAGCACGAGGAAGCGGTACGCCCATTCGCCGCCAA 240
DB 2759 GCGAATCGGAGCGGCGATACCGTAAAGCACGAGGAAGCGGTACGCCCATTCGCCGCCAA 2818
QY 241 GCTCTTCAGCAATATACAGCGGTAGCAACGCTATGCTCTGATAGCGGTCCGCGCACACCA 300
DB 2819 GCTCTTCAGCAATATACAGCGGTAGCAACGCTATGCTCTGATAGCGGTCCGCGCACACCA 2878
QY 301 GCGCGCCACAGTCGATGAATCCAGAAAGCGGCCATTTTCCACCATGATATTTTCGCAAGC 360
DB 2879 GCGCGCCACAGTCGATGAATCCAGAAAGCGGCCATTTTCCACCATGATATTTTCGCAAGC 2938
QY 361 AGGCATCGCCATGCGGTACGAGGAGATCTCGCGTCCGCGGATCGCGCTTGGCCTGG 420
DB 2939 AGGCATCGCCATGCGGTACGAGGAGATCTCGCGTCCGCGGATCGCGCTTGGCCTGG 2998
QY 421 CGAACAGTTTCGCTGCGCGAGCGCTGATGCTCTTCTGCTCCAGATCATCTTCATGACAA 480
DB 2999 CGAACAGTTTCGCTGCGCGAGCGCTGATGCTCTTCTGCTCCAGATCATCTTCATGACAA 3058
QY 481 GACCGCTTCCATCCGAGTACGTGCTCGCTGATGCGATGTTTCCTTGGTGGTGAATG 540
DB 3059 GACCGCTTCCATCCGAGTACGTGCTCGCTGATGCGATGTTTCCTTGGTGGTGAATG 3118
QY 541 GGCAGGTAGCCGATCAAGCGTATGACGCGCGCATTCGATTCGATGATGATGATCTT 600
DB 3119 GGCAGGTAGCCGATCAAGCGTATGACGCGCGCATTCGATTCGATGATGATGATCTT 3178
QY 601 TCTCGCGAGGAGCAAGGTGAGATGACGAGATCTTCCCGCGGACTTCGCCCAATAGCA 660
DB 3179 TCTCGCGAGGAGCAAGGTGAGATGACGAGATCTTCCCGCGGACTTCGCCCAATAGCA 3238
QY 661 GCGAGTCCCTTCCCGCTTCAGTGACACGTCGAGCAGACGTCGCGCAAGGACGCCGTCG 720
DB 3239 GCGAGTCCCTTCCCGCTTCAGTGACACGTCGAGCAGACGTCGCGCAAGGACGCCGTCG 3298
QY 721 TGGCAGCCAGATAGCGCGCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 3299 TGGCAGCCAGATAGCGCGCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3358
QY 781 CGGCTCTTCACAAAGACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
DB 3359 CGGCTCTTCACAAAGACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3418
QY 841 AGCAGCCGATTTGCTGTTGTCGCCAGTCATAGCCGAATAGCTCTCCACCCCAAGCGCGC 900
DB 3419 AGCAGCCGATTTGCTGTTGTCGCCAGTCATAGCCGAATAGCTCTCCACCCCAAGCGCGC 3478
QY 901 GAGAACCTTCGCTGCAATCCATCTTGTTCATCATGCGAAAGCATCTCATCTCTCTCTT 960
DB 3479 GAGAACCTTCGCTGCAATCCATCTTGTTCATCATGCGAAAGCATCTCATCTCTCTCTT 3538
QY 961 GATCAGATCTTGCATCCCTGCGCCATCAGATCCCTTGGCGGC 1001
DB 3539 GATCAGATCTTGCATCCCTGCGCCATCAGATCCCTTGGCGGC 3579

RESULT 4

US-09-042-105-16
; Sequence 16, Application US/09042105
; Patent No. 6040157
; GENERAL INFORMATION:
; APPLICANT: HU, JING-SHAN
; APPLICANT: ROSEN, CRAIG A.
; APPLICANT: CAO, LIANG
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX
; STREET: 1100 NEW YORK AVENUE
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA

```

1  ZIP: 20005
2  COMPUTER READABLE FORM:
3  MEDIUM TYPE: Floppy disk
4  COMPUTER: IBM PC compatible
5  OPERATING SYSTEM: PC-DOS/MS-DOS
6  SOFTWARE: PatentIn Release #1.0, Version #1.30
7  CURRENT APPLICATION DATA:
8  APPLICATION NUMBER: US/09/042,105
9  FILING DATE: HERewith
10 CLASSIFICATION:
11 PRIOR APPLICATION DATA:
12 APPLICATION NUMBER: US 08/207,550
13 FILING DATE: 8-MAR-1994
14 CLASSIFICATION:
15 PRIOR APPLICATION DATA:
16 APPLICATION NUMBER: US 08/465,968
17 FILING DATE: 06-JUN-1995
18 CLASSIFICATION:
19 PRIOR APPLICATION DATA:
20 APPLICATION NUMBER: TO BE ASSIGNED
21 FILING DATE: 24-DEC-1997
22 CLASSIFICATION:
23 ATTORNEY/AGENT INFORMATION:
24 NAME: ERIC K. STEFFE
25 REGISTRATION NUMBER: 36,688
26 REFERENCE/DOCKET NUMBER: 1488.1000003/EKS
27 TELECOMMUNICATION INFORMATION:
28 TELEPHONE: (202)371-2600
29 TELEFAX: (202)371-2540
30 INFORMATION FOR SEQ ID NO: 16:
31 SEQUENCE CHARACTERISTICS:
32 LENGTH: 3974 base pairs
33 TYPE: nucleic acid
34 STRANDEDNESS: both
35 TOPOLOGY: both
36 MOLECULE TYPE: DNA
37 US-09-042-105-16
38
39 Query Match 100.0%; Score 1001; DB 3; Length 3974;
40 Best Local Similarity 100.0%; Pred. No. 0;
41 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
42
43 QY 1 CGCTGGGAGATCATCCAGCCGGCTCCCGGAAACGATTCGGAGCCCAACCTTTCAATAG 60
44 Db 2579 CGCTGGGAGATCATCCAGCCGGCTCCCGGAAACGATTCGGAGCCCAACCTTTCAATAG 2638
45
46 QY 61 AAGCGGGCGGTGCAATCGAAATCTCGTGATGGCAGGTTGGGGCTCGCTTGGTCGGTCATT 120
47 Db 2639 AAGCGGGCGGTGGAATCGAAATCTCGTGATGGCAGGTTGGGGCTCGCTTGGTCGGTCATT 2698
48
49 QY 121 TCGAACCCCGAGATGCCCTCAGAAAGAACTCGTCAAGAAAGGCGATAGAGCGCATGCGCT 180
50 Db 2699 TCGAACCCCGAGATGCCCTCAGAAAGAACTCGTCAAGAAAGGCGATAGAGCGCATGCGCT 2758
51
52 QY 181 GCGAATCGGAGCGCGATACCGTAAGACAGAGGAGCGGTACGCCCATTCGCGGCCAA 240
53 Db 2759 GCGAATCGGAGCGCGGATACCGTAAGACAGAGGAGCGGTACGCCCATTCGCGGCCAA 2818
54
55 QY 241 GCTCTTCAGCAATATCACGGGTAGCCCAACGCTATGCTCTGATAGCGGTCCGCCACACCCA 300
56 Db 2819 GCTCTTCAGCAATATCACGGGTAGCCCAACGCTATGCTCTGATAGCGGTCCGCCACACCCA 2878
57
58 QY 301 GCGGCCCAAGTCGATGATCCAGAAAGGCGCATTTTCCACCATGATATTCGGCAAGC 360
59 Db 2879 GCGGCCCAAGTCGATGATCCAGAAAGGCGCATTTTCCACCATGATATTCGGCAAGC 2938
60
61 QY 361 AGGCATCGCCATGGGTCAACACAGAGATCTCTGCGCGTCGGGCGATGCGCGCTTGAGCCCTGG 420
62 Db 2939 AGGCATCGCCATGGGTCAACACAGAGATCTCTGCGCGTCGGGCGATGCGCGCTTGAGCCCTGG 2998
63
64 QY 421 CGAAACAGTTCGGTGGCGAGGCCCTTGATGCTCTTCGTCACAGATCATCTGATCGACAA 480
65 Db 2999 CGAAACAGTTCGGTGGCGAGGCCCTTGATGCTCTTCGTCACAGATCATCTGATCGACAA 3058

```



```

/ / APPLICATION NUMBER: US 60/055,561
/ / FILING DATE: 13-AUG-1997
/ / ATTORNEY/AGENT INFORMATION:
/ / NAME: STEFFE, ERIC K.
/ / REGISTRATION NUMBER: 36,688
/ / REFERENCE/POCKET NUMBER: 1488.036
/ / TELECOMMUNICATION INFORMATION:
/ / TELEPHONE: 202-371-2600
/ / TELEFAX: 202-371-2540
/ / INFORMATION FOR SEQ ID NO: 147:
/ / SEQUENCE CHARACTERISTICS:
/ / LENGTH: 3974 base pairs
/ / TYPE: nucleic acid
/ / STRANDEDNESS: both
/ / TOPOLOGY: both
/ / MOLECULE TYPE: cdna
/ / US-09-023-082A-147

```

Query Match	100.0%;	Score 1001;	DB 3;	Length 3974;
Best Local Similarity	100.0%;	Prod. 0;	Mismatches 0;	Gaps 0;
Matches 1001;	Conservative	0;	Matches 0;	Indels 0;
Qy	1	CGCTGGAGGATCATCCACCGCGCTCCGGAAACGATTCGAAAGCCCAACTTTCATAG	60	
Db	2579	CGCTGGAGGATCATCCACCGCGCTCCGGAAACGATTCGAAAGCCCAACTTTCATAG	2638	
Qy	61	AAGCGCGCGTGGAAATCGAAATCTCGTGATGCGCAGGTTGGGCGTTCGTTGGTCCGTCATT	120	
Db	2639	AAGCGCGCGTGGAAATCGAAATCTCGTGATGCGCAGGTTGGGCGTTCGTTGGTCCGTCATT	2698	
Qy	121	TCGAACCCACAGAGTCCCGCTCAGAAGAACTCTGTCAGAAGCGGATAGAGCGATGGCT	180	
Db	2699	TCGAACCCACAGAGTCCCGCTCAGAAGAACTCTGTCAGAAGCGGATAGAGCGATGGCT	2758	
Qy	181	GCGAATCGGAGCGCGCATACCGTAAAGCACAGGAACGCTCAGCCATTTCGCCGCCAA	240	
Db	2759	GCGAATCGGAGCGCGCATACCGTAAAGCACAGGAACGCTCAGCCATTTCGCCGCCAA	2818	
Qy	241	GCTCTTCAGCAATATCAAGGATAGCCAAACGCTATGTCTTGATAGCGTTCGCGCACACCCA	300	
Db	2819	GCTCTTCAGCAATATCAAGGATAGCCAAACGCTATGTCTTGATAGCGTTCGCGCACACCCA	2878	
Qy	301	GCGCGGCACAGTCGATCAATCCAGAAAGCGGCCATTTCACCATGATATTCGGCAAGC	360	
Db	2879	GCGCGGCACAGTCGATGAAATCCAGAAAGCGGCCATTTCACCATGATATTCGGCAAGC	2938	
Qy	361	AGGCATCGCCATGGGTCACGACGAGATCTCTCGCGCTCGGCGATTCGCGGCTTCGAGCTGG	420	
Db	2939	AGGCATCGCCATGGGTCACGACGAGATCTCTCGCGCTCGGCGATTCGCGGCTTCGAGCTGG	2998	
Qy	421	CGAACAGTTCCGCTGGCGGAGCCCTTGATGCTCTTCGTCCAGATCATCTGATCGACAA	480	
Db	2999	CGAACAGTTCCGCTGGCGGAGCCCTTGATGCTCTTCGTCCAGATCATCTGATCGACAA	3058	
Qy	481	GACCGGCTTCATCCGAGTACGTGCTCGCTCGATGCGATGCTTCGTCCAGATCATCTGATCGAATG	540	
Db	3059	GACCGGCTTCATCCGAGTACGTGCTCGCTCGATGCGATGCTTCGTCCAGATCATCTGATCGAATG	3118	
Qy	541	GGCAGGTAGCCGATCAAGCGTATGCAAGCCCGCGCATTTGCATCAGCCATGATGGATACTT	600	
Db	3119	GGCAGGTAGCCGATCAAGCGTATGCAAGCCCGCGCATTTGCATCAGCCATGATGGATACTT	3178	
Qy	601	TCTCGCGCAGAGCAAGGTGAGATGACAGGAGATCTTCGCCCGGACATTTCGCCCAATAGCA	660	
Db	3179	TCTCGCGCAGAGCAAGGTGAGATGACAGGAGATCTTCGCCCGGACATTTCGCCCAATAGCA	3238	
Qy	661	GCCAGTCCCTTCGCCGTTTCAAGTCAACACGTTCGAGCAGCTGCGGCAAGGAACGCCCGTCG	720	
Db	3239	GCCAGTCCCTTCGCCGTTTCAAGTCAACACGTTCGAGCAGCTGCGGCAAGGAACGCCCGTCG	3298	
Qy	721	TGCGCAGCAGATAGCCCGCTTCCTCGCTTCGACAGTTCAATTCAGGGCACCGGACAGGT	780	
Db	3299	TGCGCAGCAGATAGCCCGCTTCCTCGCTTCGACAGTTCAATTCAGGGCACCGGACAGGT	3358	

RESULT 7

US-09-044-855A-7
Sequence 7, Application US/09044855A
Patent No. 6100389
GENERAL INFORMATION:
APPLICANT: LI, HAODONG
APPLICANT: RUBEN, STEVEN M
APPLICANT: SUTTON, GRANGER G III
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A HUMAN CHEMOTACTIC
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/044,855A
FILING DATE: HERewith
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/479,126
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/424,425
FILING DATE: 21-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/05384
FILING DATE: 16-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: STEPE, ERIC K
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.0340005
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3974 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: cdNA
US-09-044-855A-7

```
Query Match      100.0%; Score 1001; DB 3; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CGCTGGAGGATCATCCAGCGCGCTCCCGAAAAAGATTCCGAAGCCCAACCTTTCATAG 60
Db |||
QY 2579 CGCTGGAGGATCATCCAGCGCGCTCCCGAAAAAGATTCCGAAGCCCAACCTTTCATAG 2638
Db |||
QY 61 AAGCGCGCGTGGAAATCGAAATCTCGTATGCGAGGTTGGCGTTCGCTTGGTTCGCTCAT 120
Db |||
QY 2639 AAGCGCGCGTGGAAATCGAAATCTCGTATGCGAGGTTGGCGTTCGCTTGGTTCGCTCAT 2698
Db |||
QY 121 TCGAACCCAGAGTCCCGCTCAGAACTCTCTCAAGAGGCGATAGAGGCGATGCCT 180
Db |||
QY 2699 TCGAACCCAGAGTCCCGCTCAGAACTCTCTCAAGAGGCGATAGAGGCGATGCCT 2758
Db |||
QY 181 GCGAATCGGAGCGGCGATACCGTAAAGCACAGAGAGCGGTACGCCATTCCGCGCA 240
Db |||
QY 2759 GCGAATCGGAGCGGCGATACCGTAAAGCACAGAGAGCGGTACGCCATTCCGCGCA 2818
Db |||
QY 241 GCTCTTCAGCAATATCAGCGGTAGCAACGCTATGTCCTGATAGCGGTCCGCCACACCA 300
Db |||
QY 2819 GCTCTTCAGCAATATCAGCGGTAGCAACGCTATGTCCTGATAGCGGTCCGCCACACCA 2878
Db |||
QY 301 GCGGCGCACAGTCGATGAATCCAGAAAAAGCGCCATTTCACACCATGATATTCGCAAGC 360
Db |||
QY 2879 GCGGCGCACAGTCGATGAATCCAGAAAAAGCGCCATTTCACACCATGATATTCGCAAGC 2938
Db |||
QY 361 AGGCATCGCCATGGGTACGACGAGATCTCTCGCGTCCGGCATGCGCGCTTTCGACCTGG 420
Db |||
QY 2939 AGGCATCGCCATGGGTACGACGAGATCTCTCGCGTCCGGCATGCGCGCTTTCGACCTGG 2998
Db |||
QY 421 CGAAGAGTTCGCTGGCGGAGCCCTGATGCTCTCTGATGCGATTCCTGATGCGAATTCGCA 480
Db |||
QY 2999 CGAAGAGTTCGCTGGCGGAGCCCTGATGCTCTCTGATGCGATTCCTGATGCGAATTCGCA 3058
Db |||
QY 481 GACCGGCTTCATCCGAGTACGCTCGCTCGATCGGATGTTTCGTTGGTGGTTCGATG 540
Db |||
QY 3059 GACCGGCTTCATCCGAGTACGCTCGCTCGATCGGATGTTTCGTTGGTGGTTCGATG 3118
Db |||
QY 541 GCGAGGTAGCCGATCAGCGTATGCGCGCGCATTCGATGCGATGATGATGATGATGATGAT 600
Db |||
QY 3119 GCGAGGTAGCCGATCAGCGTATGCGCGCGCATTCGATGCGATGATGATGATGATGATGAT 3178
Db |||
QY 601 TCTCGGAGGAGCAAGGTGAGATGACAGAGATCTCGCGCGCATTCGCGCAATGCA 660
Db |||
QY 3179 TCTCGGAGGAGCAAGGTGAGATGACAGAGATCTCGCGCGCATTCGCGCAATGCA 3238
Db |||
QY 661 GCGATGCTTCCGCTTCACTGACACAGTTCGAGCACAGCTTCGCAAGAAACGCGCTCG 720
Db |||
QY 3239 GCGATGCTTCCGCTTCACTGACACAGTTCGAGCACAGCTTCGCAAGAAACGCGCTCG 3298
Db |||
QY 721 TGGCAGGACGATAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db |||
QY 3299 TGGCAGGACGATAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3358
Db |||
QY 781 CGGTCTTGACAAAAAGAACCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db |||
QY 3359 CGGTCTTGACAAAAAGAACCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3418
Db |||
QY 841 AGCAGCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db |||
QY 3419 AGCAGCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3478
Db |||
QY 901 GAGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db |||
QY 3479 GAGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3538
Db |||
QY 961 GATCAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1001
Db |||
QY 3539 GATCAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3579
Db |||
```

RESULT 8

US-09-078-670-4

; Sequence 4, Application US/09078670

```
; Patent No. 6143498
; GENERAL INFORMATION:
; APPLICANT: OLSEN, HENRIK S.
; APPLICANT: RUBEN, STEVEN M.
; TITLE OF INVENTION: ANTIMICROBIAL PEPTIDE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/078,670
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/046,415
; FILING DATE: 14-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0930001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3974 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cdna
; US-09-078-670-4
```

```
Query Match 100.0%; Score 1001; DB 3; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGGAGGATCATCCAGCGCGCTCCCGAAAAAGATTCCGAAGCCCAACCTTTCATAG 60
Db 2579 CGCTGGAGGATCATCCAGCGCGCTCCCGAAAAAGATTCCGAAGCCCAACCTTTCATAG 2638

QY 61 AAGCGCGCGTGGAAATCGAAATCTCGTATGCGAGGTTGGCGTTCGCTTGGTTCGCTCAT 120
Db 2639 AAGCGCGCGTGGAAATCGAAATCTCGTATGCGAGGTTGGCGTTCGCTTGGTTCGCTCAT 2698

QY 121 TCGAACCCAGAGTCCCGCTCAGAACTCTCTCAAGAGGCGATAGAGGCGATGCCT 180
Db 2699 TCGAACCCAGAGTCCCGCTCAGAACTCTCTCAAGAGGCGATAGAGGCGATGCCT 2758

QY 181 GCGAATCGGAGCGGCGATACCGTAAAGCACAGAGAGCGGTACGCCATTCCGCGCA 240
Db 2759 GCGAATCGGAGCGGCGATACCGTAAAGCACAGAGAGCGGTACGCCATTCCGCGCA 2818

QY 241 GCTCTTCAGCAATATCAGCGGTAGCAACGCTATGTCCTGATAGCGGTCCGCCACACCA 300
Db 2819 GCTCTTCAGCAATATCAGCGGTAGCAACGCTATGTCCTGATAGCGGTCCGCCACACCA 2878

QY 301 GCGGCGCACAGTCGATGAATCCAGAAAAAGCGCCATTTCACACCATGATATTCGCAAGC 360
Db 2879 GCGGCGCACAGTCGATGAATCCAGAAAAAGCGCCATTTCACACCATGATATTCGCAAGC 2938

QY 361 AGGCATCGCCATGGGTACGACGAGATCTCTCGCGTCCGGCATGCGCGCTTTCGACCTGG 420
Db 2939 AGGCATCGCCATGGGTACGACGAGATCTCTCGCGTCCGGCATGCGCGCTTTCGACCTGG 2998
```

```
QY 421 CGAACAGTTCGGTGGCGGAGCCCTGATGCTCTTCGTCAGATCATCTGATGACAA 480
Db 2999 CGAACAGTTCGGTGGCGGAGCCCTGATGCTCTTCGTCAGATCATCTGATGACAA 3058
QY 481 GACCGGCTTCATCCAGTACGCTGCTCGATGGATGTTTCGTTGGTGGTGAATG 540
Db 3059 GACCGGCTTCATCCAGTACGCTGCTCGATGGATGTTTCGTTGGTGGTGAATG 3118
QY 541 GGAGGTAGCGGATCAAGCGTATGAGCGCGCGCATTCATCAGCCATGATGGATCTT 600
Db 3119 GGAGGTAGCGGATCAAGCGTATGAGCGCGCGCATTCATCAGCCATGATGGATCTT 3178
QY 601 TCTCGCAGCAGCAAGGTGATGACAGGAGATCTGCGCCCGCATTCGCGCAATAGCA 660
Db 3179 TCTCGCAGCAGCAAGGTGATGACAGGAGATCTGCGCCCGCATTCGCGCAATAGCA 3238
QY 661 GCCAGTCCCTTCCCGCTTCAGTGACAACTCGAGCAGATCGGCAAGGAGCGCCGTCG 720
Db 3239 GCCAGTCCCTTCCCGCTTCAGTGACAACTCGAGCAGATCGGCAAGGAGCGCCGTCG 3298
QY 721 TGSCCAGCAGATAGCGCGTCTGCTCTGAGTTTCTGAGTTTCTGAGGACCGGACAGGT 780
Db 3299 TGSCCAGCAGATAGCGCGTCTGCTCTGAGTTTCTGAGTTTCTGAGGACCGGACAGGT 3358
QY 781 CGGTCTTGCAAAAGAAACCGGCGCCCTGCTGACGACAGCTGCGCAAGGAGCGCGCATCAG 840
Db 3359 CGGTCTTGCAAAAGAAACCGGCGCCCTGCTGACGACAGCTGCGCAAGGAGCGCGCATCAG 3418
QY 841 AGCAGCGGATGTTCTGTTGCGCATGATGATGATGATGATGATGATGATGATGATGATG 900
Db 3419 AGCAGCGGATGTTCTGTTGCGCATGATGATGATGATGATGATGATGATGATGATGATG 3478
QY 901 GAGAACTCGTGCAATCATCTGTTCAATCATGCAAGGAGCGCGCATCTGCTGCTGCTT 960
Db 3479 GAGAACTCGTGCAATCATCTGTTCAATCATGCAAGGAGCGCGCATCTGCTGCTGCTT 3538
QY 961 GATCAGATCTGATCCCTCGCGCATCAGATCTTGGCGGC 1001
Db 3539 GATCAGATCTGATCCCTCGCGCATCAGATCTTGGCGGC 3579
```

RESULT 9

```
US-09-026-408-14
; Sequence 14, Application US/09026408
; Patent No. 6303338
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: PANCREAS-DERIVED PLASMINOGEN ACTIVATOR
; TITLE OF INVENTION: INHIBITOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/026,408
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/934,011
; FILING DATE: 15-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,056
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0300002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3974 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cdna
; US-09-026-408-14
```

```
Query Match 100.0%; Score 1001; DB 4; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGGAGGATCATCCAGCGCGCTCCGGAACCGATTCCGAGCCCAACCTTTTCATAG 60
Db 2579 CGTGGAGGATCATCCAGCGCGCTCCGGAACCGATTCCGAGCCCAACCTTTTCATAG 2638
QY 61 AAGCGCGCGTGAATCGAAATCTCTGATGGCAGGTTGGGCGTCTGCTGGTCAAT 120
Db 2639 AAGCGCGCGTGAATCGAAATCTCTGATGGCAGGTTGGGCGTCTGCTGGTCAAT 2698
QY 121 TCGAACCCCAAGATCGCGCTCAGAGAACTCGTCAAGAGGCGGATAGAGGCGGATCGGCT 180
Db 2699 TCGAACCCCAAGATCGCGCTCAGAGAACTCGTCAAGAGGCGGATAGAGGCGGATCGGCT 2758
QY 181 GCGAATCGGAGCGCGGATACCGTAAAGCAGGAGGCGGTGAGCCCATTCGCCGCCAA 240
Db 2759 GCGAATCGGAGCGCGGATACCGTAAAGCAGGAGGCGGTGAGCCCATTCGCCGCCAA 2818
QY 241 GCTCTTCAGCAATATCAGCGGTAGCCAAAGCTATGCTATGCTGATAGCGGTCCGACACCCA 300
Db 2819 GCTCTTCAGCAATATCAGCGGTAGCCAAAGCTATGCTATGCTGATAGCGGTCCGACACCCA 2878
QY 301 GCCGGCCACAGTGTGATGAATCCAGAAAGCGGCCATTTTCCACCATGATATTCGGCAAGC 360
Db 2879 GCCGGCCACAGTGTGATGAATCCAGAAAGCGGCCATTTTCCACCATGATATTCGGCAAGC 2938
QY 361 AGGCATCGCCATGGGTTCAGCAGAGATCTCGCGGTGCGGCGATCGCGCTTTCAGCGCTCG 420
Db 2939 AGGCATCGCCATGGGTTCAGCAGAGATCTCGCGGTGCGGCGATCGCGCTTTCAGCGCTCG 2998
QY 421 CGAACAGTTCGGTGGCGGAGCCCTGATGCTCTTCGTCAGATCATCTGATGACAA 480
Db 2999 CGAACAGTTCGGTGGCGGAGCCCTGATGCTCTTCGTCAGATCATCTGATGACAA 3058
QY 481 GACCGGCTTCATCCAGTACGCTGCTCGATGGATGTTTCGTTGGTGGTGAATG 540
Db 3059 GACCGGCTTCATCCAGTACGCTGCTCGATGGATGTTTCGTTGGTGGTGAATG 3118
QY 541 GGAGGTAGCGGATCAAGCGTATGAGCGCGCGCATTCATCAGCCATGATGGATCTT 600
Db 3119 GGAGGTAGCGGATCAAGCGTATGAGCGCGCGCATTCATCAGCCATGATGGATCTT 3178
QY 601 TCTCGCAGCAGCAGGTGATGACAGGAGATCTGCGCCCGCATTCGCGCAATAGCA 660
Db 3179 TCTCGCAGCAGCAGGTGATGACAGGAGATCTGCGCCCGCATTCGCGCAATAGCA 3238
QY 661 GCCAGTCCCTTCCCGCTTCAGTGACAACTCGAGCAGATCGGCAAGGAGCGCCGTCG 720
Db 3239 GCCAGTCCCTTCCCGCTTCAGTGACAACTCGAGCAGATCGGCAAGGAGCGCCGTCG 3298
QY 721 TGSCCAGCAGATAGCGCGTCTGCTCTGAGTTTCTGAGTTTCTGAGGACCGGACAGGT 780
Db 3299 TGSCCAGCAGATAGCGCGTCTGCTCTGAGTTTCTGAGTTTCTGAGGACCGGACAGGT 3358
QY 781 CGGTCTTGCAAAAGAAACCGGCGCCCTGCTGACGACAGCTGCGCAAGGAGCGCGCATCAG 840
```


Db 3359 CGGTCTTGACAAAAGAACCGGCGCCCTCGCTGACACCGCGGAACCGCGGCATCAG 3418
QY 841 ACACGCGGATCTCTGTGTGCGCGAGTCATAGCCGAATAGCCTCTCCACCCCAAGCGCGC 900
Db 3419 ACACGCGGATCTCTGTGTGCGCGAGTCATAGCCGAATAGCCTCTCCACCCCAAGCGCGC 3478
QY 901 GAGAACCTCGCTGCAATCCATCTTGTTCATATCATCGAAAGATCCTCATCTCTCTCTT 960
Db 3479 GAGAACCTCGCTGCAATCCATCTTGTTCATATCATCGAAAGATCCTCATCTCTCTCTT 3538
QY 961 GATCAGATCTTGATCCCTCGCCATCAGATCCTTGGCGGC 1001
Db 3539 GATCAGATCTTGATCCCTCGCCATCAGATCCTTGGCGGC 3579

RESULT 10

US-09-362-871-33
; Sequence 33, Application US/09362871
; Patent No. 6379923
; GENERAL INFORMATION:
; APPLICANT: DUAN, D. ROXANNE
; APPLICANT: SHILATFARD, ALI
; APPLICANT: CONAWAY, JOAN W.
; APPLICANT: CONAWAY, RONALD C.
; TITLE OF INVENTION: ELI2, A New Member of an ELL Family of
; TITLE OF INVENTION: RNA Polymerase II Elongation Factors
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SYNERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/362,871
; FILING DATE:
; CLASSIFICATION:

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/026,343
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:

NAME: GOLDSTEIN, JORGE A.
; REGISTRATION NUMBER: 29,021
; REFERENCE/DOCKET NUMBER: 1488.0880001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 372-2540

INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3974 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both

MOLECULE TYPE: cdna
; US-09-362-871-33

Query Match 100.0%; Score 1001; DB 4; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGAGGATCATCCAGCGCGGCTCCGGAAGATTCGGAAGCCCAACCTTTTCATAG 60
Db 2579 CGCTGAGGATCATCCAGCGCGGCTCCGGAAGATTCGGAAGCCCAACCTTTTCATAG 2638

QY 61 AAGCGCGCGGTGAATCGAATCTCTGTATGCGAGGTTGGGCGTGGCTTGGTGGTCAAT 120
Db 2639 AAGCGCGCGGTGAATCGAATCTCTGTATGCGAGGTTGGGCGTGGCTTGGTGGTCAAT 2698

QY 121 TCGAACCCAGAGTCCCGCTCAGAGAACTCTCTAAGAGGCGATAGAGGCGATCGCT 180
Db 2699 TCGAACCCAGAGTCCCGCTCAGAGAACTCTCTAAGAGGCGATAGAGGCGATCGCT 2758
QY 181 GCGAATCGGAGCGGCGATACCGTAAAGCAGCAGAGAGCGGTGAGCCCATTCGCGCCCAA 240
Db 2759 GCGAATCGGAGCGGCGATACCGTAAAGCAGCAGAGAGCGGTGAGCCCATTCGCGCCCAA 2818
QY 241 GCTCTTCAGCAATATCAGCGGTAGCAACGCTATGCTCTGTAGCGGTCCGCGCACACCCA 300
Db 2819 GCTCTTCAGCAATATCAGCGGTAGCAACGCTATGCTCTGTAGCGGTCCGCGCACACCCA 2878
QY 301 GCGCGCCACAGTCGATGAATCCAGAAAAGCGGCCATTTTCCACCATGATATTTCGCAAGC 360
Db 2879 GCGCGCCACAGTCGATGAATCCAGAAAAGCGGCCATTTTCCACCATGATATTTCGCAAGC 2938
QY 361 AGGCATCGCCATGGGTACAGACGAGATCCTCGCCCTCGGGCATCGCGCTTGGCCCTGG 420
Db 2939 AGGCATCGCCATGGGTACAGACGAGATCCTCGCCCTCGGGCATCGCGCTTGGCCCTGG 2998
QY 421 CGAAGATTCGCTCGCGCGAGCCCTCTGATGCTCTTCTCCAGATCATCTCTGATGACAAA 480
Db 2999 CGAAGATTCGCTCGCGCGAGCCCTCTGATGCTCTTCTCCAGATCATCTCTGATGACAAA 3058
QY 481 GACCGCTTCCATCCGAGTACGCTGCTCGATCGAGTGTTCGCTTGGTGGTCAATG 540
Db 3059 GACCGCTTCCATCCGAGTACGCTGCTCGATCGAGTGTTCGCTTGGTGGTCAATG 3118
QY 541 GGCAGGTAGCCGGATCAAGCGGTATGACCGCGCGCATTCGATCGCCATGATGATGATCTT 600
Db 3119 GGCAGGTAGCCGGATCAAGCGGTATGACCGCGCGCATTCGATCGCCATGATGATGATCTT 3178
QY 601 TCTCGGAGGAGCAAGGTGAGATGACGAGATCTTCCCGCGCATCTTCGCGCCCAATAGCA 660
Db 3179 TCTCGGAGGAGCAAGGTGAGATGACGAGATCTTCCCGCGCATCTTCGCGCCCAATAGCA 3238
QY 661 GCCAGTCCCTTCCCGCTTCAGTGACACGTCGAGCAGACGTCGCGCAAGGAGCGCCGTCG 720
Db 3239 GCCAGTCCCTTCCCGCTTCAGTGACACGTCGAGCAGACGTCGCGCAAGGAGCGCCGTCG 3298
QY 721 TGGCAGCCACGATAGCGCGCTCGCTCTCTGCTCTCAGTTCATTGAGGCGCACCGGACAGGT 780
Db 3299 TGGCAGCCACGATAGCGCGCTCGCTCTCTGCTCTCAGTTCATTGAGGCGCACCGGACAGGT 3358
QY 781 CGGTCTTGACAAAAGAACCGGCGCCCTCGCTGCGCTGACGCGCGGAAACACGCGCGCATCAG 840
Db 3359 CGGTCTTGACAAAAGAACCGGCGCCCTCGCTGCGCTGACGCGCGGAAACACGCGCGCATCAG 3418
QY 841 AGCAGCGGATTTGCTGTGTGTCGCCAGTCATAGCCGAATAGCCTCTCCACCCCAAGCGCGCG 900
Db 3419 AGCAGCGGATTTGCTGTGTGTCGCCAGTCATAGCCGAATAGCCTCTCCACCCCAAGCGCGCG 3478
QY 901 GAGAACCTCGGTGCAATCCATCTTGTTCATATCGGAAAGCATCTCTCATCTCTCTCTT 960
Db 3479 GAGAACCTCGGTGCAATCCATCTTGTTCATATCGGAAAGCATCTCTCATCTCTCTCTT 3538
QY 961 GATCAGATCTTGATCCCTCGCCATCAGATCCTTGGCGGC 1001
Db 3539 GATCAGATCTTGATCCCTCGCCATCAGATCCTTGGCGGC 3579

RESULT 11

US-09-627-154-4
; Sequence 4, Application US/09627154
; Patent No. 6420116
; GENERAL INFORMATION:
; APPLICANT: Olsen, Henrik S.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Antimicrobial Peptide
; FILE REFERENCE: 1488.0930002
; CURRENT APPLICATION NUMBER: US/09/627,154
; CURRENT FILING DATE: 2000-07-27

;; PRIOR APPLICATION NUMBER: US 09/078,670
;; PRIOR FILING DATE: 1998-05-14
;; PRIOR APPLICATION NUMBER: US 60/046,415
;; PRIOR FILING DATE: 1997-05-14
;; NUMBER OF SEQ ID NOS: 10
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 4
;; LENGTH: 3974
;; TYPE: DNA
;; ORGANISM: synthetic construct
US-09-627-154-4

Query Match 100.0%; Score 1001; DB 4; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGGAGATCATCCAGCGGGCTCCCGGAAACGATTCGGAAGCCCAACCTTTCATAG 60
Db 2579 CGTGGAGATCATCCAGCGGGCTCCCGGAAACGATTCGGAAGCCCAACCTTTCATAG 2638

QY 61 AAGCGCGCGTGAATCGAAATCTCGTATGCGAGGTGGGGCTGGCTTGGTGGTGGTATT 120
Db 2639 AAGCGCGCGTGAATCGAAATCTCGTATGCGAGGTGGGGCTGGCTTGGTGGTGGTATT 2698

QY 121 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGCGATAGAGGCGATGCGCT 180
Db 2699 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGCGATAGAGGCGATGCGCT 2758

QY 181 GCGAATCGGAGCGGCGATACCGTAAAGCAGAGAGCGGTGACGCCATTCGCGCAAC 240
Db 2759 GCGAATCGGAGCGGCGATACCGTAAAGCAGAGAGCGGTGACGCCATTCGCGCAAC 2818

QY 241 GCTCTTCAGCAATATCAGCGGTAGCAACGCTATCTCTGATAGCGGTCCGCCACACCA 300
Db 2819 GCTCTTCAGCAATATCAGCGGTAGCAACGCTATCTCTGATAGCGGTCCGCCACACCA 300

QY 301 GCGGCGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
Db 2879 GCGGCGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2938

QY 361 AGGCGCGCGTGAATCGAAATCTCGTATGCGAGGTGGGGCTGGCTTGGTGGTGGTATT 420
Db 2939 AGGCGCGCGTGAATCGAAATCTCGTATGCGAGGTGGGGCTGGCTTGGTGGTGGTATT 2998

QY 421 CGAACAGTTTCGCTGGCGAGCCCTCGATGCTCTTCGTCAGAGTATCTCTGATGCGCA 480
Db 2999 CGAACAGTTTCGCTGGCGAGCCCTCGATGCTCTTCGTCAGAGTATCTCTGATGCGCA 3058

QY 481 GACCGGCTTCATCGAGTACGTGCTGCTCGATGCGATGCTTGGTGGTGGTGGTATT 540
Db 3059 GACCGGCTTCATCGAGTACGTGCTGCTCGATGCGATGCTTGGTGGTGGTGGTATT 3118

QY 541 GCGAGGTAGCGGATCAAGCGTATGCGCGCGCATTCGATCAGCCATGATGATGATGAT 600
Db 3119 GCGAGGTAGCGGATCAAGCGTATGCGCGCGCATTCGATCAGCCATGATGATGATGAT 3178

QY 601 TCTCGGAGGAGCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Db 3179 TCTCGGAGGAGCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3238

QY 661 GCGAGTCCCTTCCTCGCTTCAGTGAACAAGTTCAGAGTTCGAGTTCGAGTTCGAGT 720
Db 3239 GCGAGTCCCTTCCTCGCTTCAGTGAACAAGTTCAGAGTTCGAGTTCGAGTTCGAGT 3298

QY 721 TGGCAGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
Db 3299 TGGCAGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3358

QY 781 CGGCTTCGACAAAAGAACCGCGCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 3359 CGGCTTCGACAAAAGAACCGCGCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3418

QY 841 AGCAGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900

Db 3419 AGCAGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3478
QY 901 GAGAACCTGCGTGCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 960
Db 3479 GAGAACCTGCGTGCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 3538

QY 961 GATCAGATCTTGATCCCTCGGCGATCAGATCCTTGCGCGC 1001
Db 3539 GATCAGATCTTGATCCCTCGGCGATCAGATCCTTGCGCGC 3579

RESULT 12
US-09-027-287-50
; Sequence 50, Application US/09027287A
; Patent No. 6479254
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ruben, Steven M.
; APPLICANT: Ullrich, Stephen
; TITLE OF INVENTION: Apoptosis Inducing Molecule II
; FILE REFERENCE: 1488.065004
; CURRENT APPLICATION NUMBER: US/09/027,287A
; CURRENT FILING DATE: 1998-02-20
; EARLIER APPLICATION NUMBER: US 09/003,886
; EARLIER FILING DATE: 1998-01-07
; EARLIER APPLICATION NUMBER: US 08/822,953
; EARLIER FILING DATE: 1997-03-21
; EARLIER APPLICATION NUMBER: US 60/030,157
; EARLIER FILING DATE: 1996-10-31
; EARLIER APPLICATION NUMBER: US 60/013,923
; EARLIER FILING DATE: 1996-03-22
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50
; LENGTH: 3974
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: vector
US-09-027-287-50

Query Match 100.0%; Score 1001; DB 4; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGGAGATCATCCAGCGGGCTCCCGGAAACGATTCGGAAGCCCAACCTTTCATAG 60
Db 2579 CGTGGAGATCATCCAGCGGGCTCCCGGAAACGATTCGGAAGCCCAACCTTTCATAG 2638

QY 61 AAGCGCGCGTGAATCGAAATCTCGTATGCGAGGTGGGGCTGGCTTGGTGGTGGTATT 120
Db 2639 AAGCGCGCGTGAATCGAAATCTCGTATGCGAGGTGGGGCTGGCTTGGTGGTGGTATT 2698

QY 121 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGCGATAGAGGCGATGCGCT 180
Db 2699 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGCGATAGAGGCGATGCGCT 2758

QY 181 GCGAATCGGAGCGGCGATACCGTAAAGCAGAGAGCGGTGACGCCATTCGCGCAAC 240
Db 2759 GCGAATCGGAGCGGCGATACCGTAAAGCAGAGAGCGGTGACGCCATTCGCGCAAC 2818

QY 241 GCTCTTCAGCAATATCAGCGGTAGCAACGCTATCTCTGATAGCGGTCCGCCACACCA 300
Db 2819 GCTCTTCAGCAATATCAGCGGTAGCAACGCTATCTCTGATAGCGGTCCGCCACACCA 2878

QY 301 GCGGCGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
Db 2879 GCGGCGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2938

QY 361 AGGCGCGCGTGAATCGAAATCTCGTATGCGAGGTGGGGCTGGCTTGGTGGTGGTATT 420

Db 2939 AGGCATCGCCATGGGTCAACGACGAGATCTCTCGCGCTCGGCGATCGCGCTTTGAGCCCTGG 2998
QY 421 CGACAGCTTGGGCTGGGCGAGCCCTGATGCTCTTCGTCAGATCATCTCTGATCGACAA 480
Db 2999 CGACAGTTCGGCTGGGCGAGCCCTGATGCTCTTCGTCAGATCATCTCTGATCGACAA 3058
QY 481 GACCGGCTTCATCCGAGTACGTCGCTCGCTCGATGCGATGTTTCGCTTGGTGGTGAATG 540
Db 3059 GACCGGCTTCATCCGAGTACGTCGCTCGCTCGATGCGATGTTTCGCTTGGTGGTGAATG 3118
QY 541 GGCAGGTAGCGGATCAAGCGTATGACGCGCGCGATTCATCATCGACCATGATGATCTT 600
Db 3119 GGCAGGTAGCGGATCAAGCGTATGACGCGCGCGATTCATCATCGACCATGATGATCTT 3178
QY 601 TCTCGGAGGACGAGTGAATGACGAGGATCTCTCGCCCGGCACTTCGCCCAATAGCA 660
Db 3179 TCTCGGAGGACGAGTGAATGACGAGGATCTCTCGCCCGGCACTTCGCCCAATAGCA 3238
QY 661 GCGAGTCCCTTCGCTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 720
Db 3239 GCGAGTCCCTTCGCTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 3298
QY 721 TGGCCAGCCAGTATGAGCGGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 780
Db 3299 TGGCCAGCCAGTATGAGCGGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 3358
QY 781 CGGTCTTTCGACAAAGAACCGGCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 840
Db 3359 CGGTCTTTCGACAAAGAACCGGCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 3418
QY 841 AGCAGCGATTTGCTGTGTGTCGCGAGTCAAGCGGATAGCTCTCCAGCCAGCGGCGCG 900
Db 3419 AGCAGCGATTTGCTGTGTGTCGCGAGTCAAGCGGATAGCTCTCCAGCCAGCGGCGCG 3478
QY 901 GAGAACCTGGTGAATCCATCTGTTCAATCATGCGAAACGATCTCTCTCTCTCTCT 960
Db 3479 GAGAACCTGGTGAATCCATCTGTTCAATCATGCGAAACGATCTCTCTCTCTCTCTCT 3538
QY 961 GATCAGATCTTGTATCCCTTCGCGCATCATGATCTTTCGCGGC 1001
Db 3539 GATCAGATCTTGTATCCCTTCGCGCATCATGATCTTTCGCGGC 3579

RESULT 13

US-09-437-602-4
; Sequence 4, Application US/09437602
; Patent No. 6495128
; GENERAL INFORMATION:
; APPLICANT: Salcedo, Theodora W.
; APPLICANT: Patel, Vikram, P.
; APPLICANT: Nibbs, Robert J.B.
; APPLICANT: Graham, Gerard J.
; TITLE OF INVENTION: CHEMOKINE BETA-7
; FILE REFERENCE: 1488.033000F
; CURRENT APPLICATION NUMBER: US/09/437,602
; PRIOR FILING DATE: 1999-11-10
; PRIOR APPLICATION NUMBER: US 60/107,801
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 4
; LENGTH: 3974
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-437-602-4

Query Match 100.0%; Score 1001; DB 4; Length 3974;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGGAGGATCATCCAGCGGGGTCCCGGAAACGATTCGGAAGCCCAACCTTTCATAG 60
Db 2579 CGCTGGAGGATCATCCAGCGGGGTCCCGGAAACGATTCGGAAGCCCAACCTTTCATAG 2638

QY 61 AAGCGCGGCTGGAATCGAATCTCGTATGCGAGCTTGGCGCTCGCTTGGTTCGGTATT 120
Db 2639 AAGCGCGGCTGGAATCGAATCTCGTATGCGAGCTTGGCGCTCGCTTGGTTCGGTATT 2698
QY 121 TCGAACCCAGAGTCCCGCTCAGAAAGACTCGTCAAGAGGCGATAGAGGCGATGCGCT 180
Db 2699 TCGAACCCAGAGTCCCGCTCAGAAAGACTCGTCAAGAGGCGATAGAGGCGATGCGCT 2758
QY 181 GCGAATCGGAGCGGCGATACCGTAAAGCACAGAGAGCGGTGAGCCCATTCGCGCCAA 240
Db 2759 GCGAATCGGAGCGGCGATACCGTAAAGCACAGAGAGCGGTGAGCCCATTCGCGCCAA 2819
QY 241 GCTCTTTCAGCAATATCAGCGGTAGCAACGCTATGCTCTGATAGCGGTCCGCCACACCA 300
Db 2819 GCTCTTTCAGCAATATCAGCGGTAGCAACGCTATGCTCTGATAGCGGTCCGCCACACCA 2878
QY 301 GCGGCGCACAGTTCGATCCAGAAAGCGGCCATTTCCACCATGATATTCGSCAAGC 360
Db 2879 GCGGCGCACAGTTCGATCCAGAAAGCGGCCATTTCCACCATGATATTCGSCAAGC 2938
QY 361 AGGCATCGCCATGGGTACGACGAGATCCTCGCCGTCGCGCATGCGCCTTGAAGCTGG 420
Db 2939 AGGCATCGCCATGGGTACGACGAGATCCTCGCCGTCGCGCATGCGCCTTGAAGCTGG 2998
QY 421 CGAACAGTTCGGTGGCGGAGCCCTGATGCTCTTCGTCAGATCATCTGATCGACAA 480
Db 2999 CGAACAGTTCGGTGGCGGAGCCCTGATGCTCTTCGTCAGATCATCTGATCGACAA 3058
QY 481 GACCGGCTTCCATTCGAGTACGTCGCTCGCTCGATCGATGTTTCGCTTGGTGGTGAATG 540
Db 3059 GACCGGCTTCCATTCGAGTACGTCGCTCGCTCGATCGATGTTTCGCTTGGTGGTGAATG 3118
QY 541 GCGAGTTCGGTTCGCTTCAGTTCAGCGCGCGCATTCGATCGACCATGATGATGATCTT 600
Db 3119 GCGAGTTCGGTTCGCTTCAGTTCAGCGCGCGCATTCGATCGACCATGATGATGATCTT 3178
QY 601 TCTCGGAGGACGAGTGAATGACGAGATCCTCGCCCGGCACTTCGCCCAATAGCA 660
Db 3179 TCTCGGAGGACGAGTGAATGACGAGATCCTCGCCCGGCACTTCGCCCAATAGCA 3238
QY 661 GCGAGTTCCTTCGCTTCAGTTCAGCGCGCGCATTCGCGCACAGCTGCGCAAGAGCGCCGTCG 720
Db 3239 GCGAGTTCCTTCGCTTCAGTTCAGCGCGCGCATTCGCGCACAGCTGCGCAAGAGCGCCGTCG 3298
QY 721 TGGCCAGCCAGTATGAGCGGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 780
Db 3299 TGGCCAGCCAGTATGAGCGGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 3358
QY 781 CGGTCTTTCGACAAAGAACCGGCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 840
Db 3359 CGGTCTTTCGACAAAGAACCGGCGCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 3418
QY 841 AGCAGCGGATTTGCTGTGTGTCGCGAGTCAAGCGGATAGCTCTTCACCCAGCGGCGCG 900
Db 3419 AGCAGCGGATTTGCTGTGTGTCGCGAGTCAAGCGGATAGCTCTTCACCCAGCGGCGCG 3478
QY 901 GAGAACCTGGTGAATCCATCTGTTCAATCATGCGAAACGATCTCTCTCTCTCTCT 960
Db 3479 GAGAACCTGGTGAATCCATCTGTTCAATCATGCGAAACGATCTCTCTCTCTCTCTCT 3538
QY 961 GATCAGATCTTGTATCCCTTCGCGCATCATGATCTTTCGCGGC 1001
Db 3539 GATCAGATCTTGTATCCCTTCGCGCATCATGATCTTTCGCGGC 3579

RESULT 14

US-09-252-656B-50

; Sequence 50, Application US/09252656B

; Patent No. 6495520

; GENERAL INFORMATION:

; APPLICANT: Ebner, Reinhard

; APPLICANT: Yu, Guo-Liang

APPLICANT: Ruben, Steven M.
APPLICANT: Zhang, Jun
APPLICANT: Ullrich, Stephen
APPLICANT: Zhai, Yifan
TITLE OF INVENTION: Apoptosis Inducing Molecule II and Methods of Use
FILE REFERENCE: 1488.065006
CURRENT APPLICATION NUMBER: US/09/252,656B
CURRENT FILING DATE: 1999-02-19
PRIORITY APPLICATION NUMBER: US 60/075,409
PRIORITY FILING DATE: 1998-02-20
PRIORITY APPLICATION NUMBER: US 09/027,287
PRIORITY FILING DATE: 1998-02-20
PRIORITY APPLICATION NUMBER: US 09/003,886
PRIORITY FILING DATE: 1998-01-07
PRIORITY APPLICATION NUMBER: US 08/822,953
PRIORITY FILING DATE: 1997-03-21
PRIORITY APPLICATION NUMBER: US 60/013,923
PRIORITY FILING DATE: 1996-03-22
PRIORITY APPLICATION NUMBER: US 60/030,157
PRIORITY FILING DATE: 1996-10-31
NUMBER OF SEQ ID NOS: 61
SOFTWARE: PatentIn version 3.1
SEQ ID NO 50
LENGTH: 3974
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: pHE4-5 expression vector
US-09-252-656B-50

Query Match 100.0%; Score 100.; DB 4; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGAGGATCATCCAGCGCGCTCCGGAAGAGTTCGGAAGCCCACTTTTCATAG 60
DB 2579 CGCTGAGGATCATCCAGCGCGCTCCGGAAGAGTTCGGAAGCCCACTTTTCATAG 2638

QY 61 AAGCGCGGTTGGAATCGAAATCTCGTATGCGAGTGGCGCTGCTTGGTTCGGTCAAT 120
DB 2639 AAGCGCGGTTGGAATCGAAATCTCGTATGCGAGTGGCGCTGCTTGGTTCGGTCAAT 2698

QY 121 TCGAACCCAGAGTCCCGCTCAGAGACTGTCAGAGCGGATAGAGCGATGCGCT 180
DB 2699 TCGAACCCAGAGTCCCGCTCAGAGACTGTCAGAGCGGATAGAGCGATGCGCT 2758

QY 181 GCGAATCGGGAGCGCGATACCGTAAAGCAGAGAGCGGTGAGCCATTGCGCGCAA 240
DB 2759 GCGAATCGGGAGCGCGATACCGTAAAGCAGAGAGCGGTGAGCCATTGCGCGCAA 2818

QY 241 GCTCTTCAGCAATATCAGGAGTACGCAACGCTATGCTGATAGCGGTGCGGCACACCA 300
DB 2819 GCTCTTCAGCAATATCAGGAGTACGCAACGCTATGCTGATAGCGGTGCGGCACACCA 2878

QY 301 GCGGCGCCAGTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATATTCGGCAAGC 360
DB 2879 GCGGCGCCAGTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATATTCGGCAAGC 2938

QY 361 AGGCAATCGGATGCGGTACGACGAGATCTCTCCGCTCGGGCATCGCGCTTCGAGCCTGG 420
DB 2939 AGGCAATCGGATGCGGTACGACGAGATCTCTCCGCTCGGGCATCGCGCTTCGAGCCTGG 2998

QY 421 CGAACAGTTCGGCTGGCGGAGCCCTGATGCTCTGTCAGATCATCTGATCGACAA 480
DB 2999 CGAACAGTTCGGCTGGCGGAGCCCTGATGCTCTGTCAGATCATCTGATCGACAA 3058

QY 481 GACCGGCTTCATCCGAGTACGTCTCGTCTGATGCGATGTTTCGCTTGGTGGTTCGAATG 540
DB 3059 GACCGGCTTCATCCGAGTACGTCTCGTCTGATGCGATGTTTCGCTTGGTGGTTCGAATG 3118

QY 541 GGCAGGTAGCGGATCAGCGTATGACGCGCCGCTGATCAGCCATGATGATCTT 600
DB 3119 GGCAGGTAGCGGATCAGCGTATGACGCGCCGCTGATCAGCCATGATGATCTT 3178

QY 601 TCTCGGAGGAGCAAGTTCAGATGACAGGAGATCTTCCCGCGCACTTCGCCCAATAGCA 660
DB 3179 TCTCGGAGGAGCAAGTTCAGATGACAGGAGATCTTCCCGCGCACTTCGCCCAATAGCA 3238

QY 661 GCGAGTCCCTTCGCGCTTCAGTGACAACTGAGCAGCAGTTCAGGAGCAAGCAAGCGCGTGC 720
DB 3239 GCGAGTCCCTTCGCGCTTCAGTGACAACTGAGCAGCAGTTCAGGAGCAAGCAAGCGCGTGC 3298

QY 721 TGGCCAGCCAGCAGTACCGCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 3299 TGGCCAGCCAGCAGTACCGCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3358

QY 781 CGGTCTTGACAAAAGAACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
DB 3359 CGGTCTTGACAAAAGAACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3418

QY 841 AGCAGCGGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
DB 3419 AGCAGCGGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3478

QY 901 GAGAACCTCGTCAATCCATCTTGTTCATCATCGAAACGATCTCTCATCTCTCTT 960
DB 3479 GAGAACCTCGTCAATCCATCTTGTTCATCATCGAAACGATCTCTCATCTCTCTT 3538

QY 961 GATCAGATCTTGCATCCCTGCGCCATCAGATCTTGGCGGC 1001
DB 3539 GATCAGATCTTGCATCCCTGCGCCATCAGATCTTGGCGGC 3579

RESULT 15
US-09-044-796A-10
Sequence 10; Application US/09044796A
Patent No. 6194168
GENERAL INFORMATION:
APPLICANT: Gentz, Reiner L.
APPLICANT: Coleman, Timothy A.
TITLE OF INVENTION: Expression Control Sequences
FILE REFERENCE: PV447
CURRENT APPLICATION NUMBER: US/09/044,796A
CURRENT FILING DATE: 1998-03-20
PRIORITY APPLICATION NUMBER: 08/941,020
PRIORITY FILING DATE: 1997-09-30
PRIORITY APPLICATION NUMBER: 09/044,796
PRIORITY FILING DATE: 1998-03-20
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 3984
TYPE: DNA
ORGANISM: Homo sapiens
US-09-044-796A-10

Query Match 100.0%; Score 100.; DB 3; Length 3984;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGAGGATCATCCAGCGCGCTCCGGAAGAGTTCGGAAGCCCACTTTTCATAG 60
DB 2701 CGCTGAGGATCATCCAGCGCGCTCCGGAAGAGTTCGGAAGCCCACTTTTCATAG 2760

QY 61 AAGCGCGGTTGGAATCGAAATCTCGTATGCGAGTGGCGCTGCTTGGTTCGGTCAAT 120
DB 2761 AAGCGCGGTTGGAATCGAAATCTCGTATGCGAGTGGCGCTGCTTGGTTCGGTCAAT 2820

QY 121 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGCGGATGCGCT 180
DB 2821 TCGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGCGGATGCGCT 2880

QY 181 GCGAATCGGGAGCGCGATACCGTAAAGCAGAGAGCGGTGAGCCATTTCGCGCCAA 240
DB 2881 GCGAATCGGGAGCGCGATACCGTAAAGCAGAGAGCGGTGAGCCATTTCGCGCCAA 2940

Qy 241 GCTCTTACGAATATACGGGTAGCCACGCTATGTCTCTGATAGCGTCCGCCACACCCA 300
Db |||||
2941 GCTCTTACGAATATACGGGTAGCCACGCTATGTCTCTGATAGCGTCCGCCACACCCA 3000
Qy 301 GCCGCCACAGTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATATTCGGCAAGC 360
Db |||||
3001 GCCGCCACAGTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATATTCGGCAAGC 3060
Qy 361 AGGCATGCCATGGTACAGCAGAGATCCTCGCGTGGGCATGCGCGCTTGAGCCTGG 420
Db |||||
3061 AGGCATGCCATGGTACAGCAGAGATCCTCGCGTGGGCATGCGCGCTTGAGCCTGG 3120
Qy 421 CGAACAGTTCGGTGGCGCGAGCCCTGATCTCTTCTGTCAGATCATCTGATCGACAA 480
Db |||||
3121 CGAACAGTTCGGTGGCGCGAGCCCTGATCTCTTCTGTCAGATCATCTGATCGACAA 3180
Qy 481 GACCGCTTCCATCGAGTAGCTGCTGCTCGATGCGATGTTTCGCTTGGTGGTGAATG 540
Db |||||
3181 GACCGCTTCCATCGAGTAGCTGCTGCTCGATGCGATGTTTCGCTTGGTGGTGAATG 3240
Qy 541 GGCAGGTAGCCGATCAAGCGTATGCAGCGCGCGCATTTGATCAGCCATGATGATACTT 600
Db |||||
3241 GGCAGGTAGCCGATCAAGCGTATGCAGCGCGCGCATTTGATCAGCCATGATGATACTT 3300
Qy 601 TCTCGGACGAGCAAGGTGAGATCACAGGAGATCCTCCCGGCACTTCGCCCAATAGCA 660
Db |||||
3301 TCTCGGACGAGCAAGGTGAGATCACAGGAGATCCTCCCGGCACTTCGCCCAATAGCA 3360
Qy 661 GCCAGTCCCTTCCGCTTCAAGTACAGCTGCGAGCAGCTGCGCAGGAAACGCCGTCG 720
Db |||||
3361 GCCAGTCCCTTCCGCTTCAAGTACAGCTGCGAGCAGCTGCGCAGGAAACGCCGTCG 3420
Qy 721 TGGCCAGCCAGATAGCCGCGCTGCTCTGTCAGTTCATTACAGGCGACCGGACAGGT 780
Db |||||
3421 TGGCCAGCCAGATAGCCGCGCTGCTCTGTCAGTTCATTACAGGCGACCGGACAGGT 3480
Qy 781 CGGTCTTGCACAAAGAAACCGGGCGCCCTGCGCTGACCGGGAACAAGCGGCAATCAG 840
Db |||||
3481 CGGTCTTGCACAAAGAAACCGGGCGCCCTGCGCTGACCGGGAACAAGCGGCAATCAG 3540
Qy 841 AGCAGCCGATTTCTGTGTGCCAGTCATAGCCGATAGCTCTCCACCCAGCGGCGG 900
Db |||||
3541 AGCAGCCGATTTCTGTGTGCCAGTCATAGCCGATAGCTCTCCACCCAGCGGCGG 3600
Qy 901 GAGAACCTGCGTGAATCCATCTTGTTCATATCGGAAACGATCCTCTCTCTCTT 960
Db |||||
3601 GAGAACCTGCGTGAATCCATCTTGTTCATATCGGAAACGATCCTCTCTCTCTT 3660
Qy 961 GATCAGATCTTGATCCCTGCGCATCAGATCCTTGGCGGC 1001
Db |||||
3661 GATCAGATCTTGATCCCTGCGCATCAGATCCTTGGCGGC 3701

Search completed: February 15, 2004, 19:10:18
Job time : 77.6147 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 16:21:35 , Search time 420.484 Seconds
(without alignments)
8769.219 Million cell updates/sec

Title: US-09-921-143-36_COPY_4000_5000

Perfect score: 1001

Sequence: 1 cgcggaggatccagcc.....gcatcagatccttgcgcg 1001

Scoring table: OULGO_NUC

Gapex 60.0 , Gapex 60.0

Searched: 2449703 seqs, 1841816367 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
C 1	1001	100.0	2406	13	US-10-307-138-14
C 2	1001	100.0	2772	13	US-10-307-138-15
C 3	1001	100.0	3974	9	US-09-027-287-50
4	1001	100.0	3974	9	US-09-252-656B-50
5	1001	100.0	3974	10	US-09-972-912-8
6	1001	100.0	3974	10	US-09-902-684-14
7	1001	100.0	3974	11	US-09-935-726-16
8	1001	100.0	3974	11	US-09-105-470-25
9	1001	100.0	3974	11	US-09-345-373-147
10	1001	100.0	3974	11	US-09-373-658-12
11	1001	100.0	3974	12	US-09-983-687-12
12	1001	100.0	3974	12	US-10-375-680-50
13	1001	100.0	3974	12	US-10-614-275-25
14	1001	100.0	3974	13	US-10-260-270-4
15	1001	100.0	3974	13	US-10-197-844-9
					Sequence 14, Appl
					Sequence 15, Appl
					Sequence 50, Appl
					Sequence 8, Appl
					Sequence 14, Appl
					Sequence 15, Appl
					Sequence 25, Appl
					Sequence 147, App
					Sequence 12, Appl
					Sequence 50, Appl
					Sequence 25, Appl
					Sequence 4, Appl

16	1001	100.0	3974	13	US-10-120-398-16	Sequence 16, Appl
17	1001	100.0	3974	13	US-10-307-138-16	Sequence 16, Appl
18	1001	100.0	3974	13	US-10-120-414-16	Sequence 16, Appl
19	1001	100.0	3974	13	US-10-035-212-147	Sequence 147, App
20	1001	100.0	3974	13	US-09-921-143-16	Sequence 16, Appl
21	1001	100.0	3974	14	US-10-028-780-33	Sequence 33, Appl
22	1001	100.0	3974	14	US-10-141-965-7	Sequence 7, Appl
23	1001	100.0	3974	15	US-10-084-488-16	Sequence 16, Appl
24	1001	100.0	3974	16	US-10-075-446-147	Sequence 147, App
25	1001	100.0	4208	15	US-10-261-950-37	Sequence 37, Appl
26	1001	100.0	4256	15	US-10-054-967-21	Sequence 21, Appl
C 27	1001	100.0	4323	13	US-10-307-138-4	Sequence 4, Appl
28	1001	100.0	4328	12	US-10-445-642-1	Sequence 1, Appl
29	1001	100.0	4328	13	US-10-278-751-2	Sequence 2, Appl
C 30	1001	100.0	4800	13	US-10-322-360-1	Sequence 1, Appl
31	1001	100.0	4818	13	US-10-278-751-4	Sequence 4, Appl
C 32	1001	100.0	4900	13	US-10-307-138-3	Sequence 3, Appl
33	1001	100.0	5107	13	US-10-278-751-3	Sequence 3, Appl
C 34	1001	100.0	5130	10	US-09-897-006-9	Sequence 9, Appl
C 35	1001	100.0	5130	11	US-09-897-511A-9	Sequence 9, Appl
C 36	1001	100.0	5156	13	US-10-307-138-12	Sequence 12, Appl
37	1001	100.0	5283	13	US-09-921-143-36	Sequence 36, Appl
38	1001	100.0	5701	11	US-03-764-891-5553	Sequence 5553, Ap
39	1001	100.0	5701	11	US-03-764-891-5553	Sequence 5553, Ap
C 40	1001	100.0	6087	13	US-10-307-138-13	Sequence 13, Appl
41	1001	100.0	6514	14	US-10-090-983-1	Sequence 1, Appl
C 42	1001	100.0	7635	13	US-10-054-536-25	Sequence 25, Appl
C 43	1001	100.0	8068	13	US-10-159-006-27	Sequence 27, Appl
C 44	1001	100.0	8068	13	US-10-159-006-35	Sequence 35, Appl
45	1001	100.0	8349	13	US-10-198-478-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1

US-10-307-138-14/c
; Sequence 14, Application US/10307138
; Publication No. US20030162267A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroeder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Habernauer, Gregor
; APPLICANT: Haefner, Stefan
; APPLICANT: Klopptogge, Corinna
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING REGULATORY
; FILE REFERENCE: BGI-123CP2
; CURRENT APPLICATION NUMBER: US/10307138
; CURRENT FILING DATE: 2002-11-23
; PRIOR APPLICATION NUMBER: 09/602874
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/422618
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142690
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: 60/151251
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930476.9
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931419.5
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932122.1
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; Remaining Prior Application data removed - See File Wrapper or PALM.

```

; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 2406
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: vector
; OTHER INFORMATION: PINTTEGRATIV
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (457)..(1248)
; OTHER INFORMATION: kanamycin resistance gene from Tn5
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1515)..(2375)
; OTHER INFORMATION: Ori from pMB for replication in E.coli
; US-10-307-138-14

Query Match      100.0%; Score 1001; DB 13; Length 2406;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCTGAGGATCATCCAGCGCGTCCCGGAAACGATTCGAAAGCCCAACCTTTTCATAG 60
DB 1390 CGCTGAGGATCATCCAGCGCGTCCCGGAAACGATTCGAAAGCCCAACCTTTTCATAG 1331
QY 61 AAGCGGGCGTGAATCGAATCTCGTATGCGAGGTTGGCGTGGCTTGGTGGTCAATT 120
DB 1330 AAGCGGGCGTGAATCGAATCTCGTATGCGAGGTTGGCGTGGCTTGGTGGTCAATT 1271
QY 121 TCGAACCCCGAGAGTCCCGCTCAGAGAACTCGTCAAGAACGCGATAGAGCGCATGCGCT 180
DB 1270 TCGAACCCCGAGAGTCCCGCTCAGAGAACTCGTCAAGAACGCGATAGAGCGCATGCGCT 1211
QY 181 GCGAATCGGAGCGCGATACGTTAAACACGAGAGCGGTACGCCATTCGCCGCCAA 240
DB 1210 GCGAATCGGAGCGCGATACGTTAAACACGAGAGCGGTACGCCATTCGCCGCCAA 1151
QY 241 GCTCTTCAGCAATATACCGGTAGCAACGATGTCTGTCTGATAGCGTCCGCCACACCA 300
DB 1150 GCTCTTCAGCAATATACCGGTAGCAACGATGTCTGTCTGATAGCGTCCGCCACACCA 1091
QY 301 GCGGCCACAGTCGATGAATCCAGAAACGCGCCATTTCCACCATGATATTCGCGAAGC 360
DB 1090 GCGGCCACAGTCGATGAATCCAGAAACGCGCCATTTCCACCATGATATTCGCGAAGC 1031
QY 361 AGGCATCGCATGGTCAACGAGATCTCGCCGTCGGCATCGCGCGCTTGAGCTGG 420
DB 1030 AGGCATCGCATGGTCAACGAGATCTCGCCGTCGGCATCGCGCGCTTGAGCTGG 971
QY 421 CGAACAGTTCGGCTGGCGGAGCCCTGATGCTCTTGTCTCAGATCATCTGTATCGACAA 480
DB 970 CGAACAGTTCGGCTGGCGGAGCCCTGATGCTCTTGTCTCAGATCATCTGTATCGACAA 911
QY 481 GACCGGCTTCATCCGAGTAGTGTCTGCTGATCGAGTTTCGTTGGTGGTGAATG 540
DB 910 GACCGGCTTCATCCGAGTAGTGTCTGCTGATCGAGTTTCGTTGGTGGTGAATG 851
QY 541 GCGAGTAGCCCGATCAAGCGTATGACGCGCGCATTTGATCAGCCATGATGATGACTT 600
DB 850 GCGAGTAGCCCGATCAAGCGTATGACGCGCGCATTTGATCAGCCATGATGATGACTT 791
QY 601 TCTCGGAGGACGAGGTGATGACAGGAGATCCTGCCCGGCACTTCGCCCAATAGCA 660
DB 790 TCTCGGAGGACGAGGTGATGACAGGAGATCCTGCCCGGCACTTCGCCCAATAGCA 731
QY 661 GCGATCCCTTCCCGCTTCAGTCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 720
DB 730 GCGATCCCTTCCCGCTTCAGTCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 671
QY 721 TGGCCAGCCAGATAGCGCGCTCGCTCGTCTGATGATTCATTCAGGCGCACCGGACAGGT 780
DB 780 TGGCCAGCCAGATAGCGCGCTCGCTCGTCTGATGATTCATTCAGGCGCACCGGACAGGT 780

; US-10-307-138-15/c
; Sequence 15, Application US/10307138
; Publication No. US20030162267A1
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroeder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauser, Gregor
; APPLICANT: Haefner, Stefan
; APPLICANT: Kloppe, Corinna
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING REGULATORY
; TITLE OF INVENTION: PROTEINS INVOLVED IN THE PRODUCTION OF METHIONINE
; FILE REFERENCE: EGI-123CP2
; CURRENT APPLICATION NUMBER: US/10/307,138
; CURRENT FILING DATE: 2002-11-29
; PRIOR APPLICATION NUMBER: 09/602874
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/422618
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142690
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: 60/151251
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930476.9
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931419.5
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932122.1
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 2772
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: vector
; OTHER INFORMATION: PINTTEGRATIV deitax02910 for knock-out of
; OTHER INFORMATION: positive regulator of methionine biosynthesis
; OTHER INFORMATION: rx02910
; NAME/KEY: misc feature
; LOCATION: (457)..(1248)
; OTHER INFORMATION: kanamycin resistance gene form Tn5
; FEATURE:

```



```

; NAME/KEY: misc_feature
; LOCATION: (1515)..(2375)
; OTHER INFORMATION: Ori from pMB for replication in E.coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(61)
; OTHER INFORMATION: fragment of rx02910 (cont. from bp 2772 in
; OTHER INFORMATION: circular plasmid)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3407)..(2772)
; OTHER INFORMATION: fragment of rx02910 (1st part to be con. from bp
; OTHER INFORMATION: 1 to 61 in circular plasmid)
US-107-138-15

```

Query Match	100.0%;	Score 1001;	DB 13;	Length 2772;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1001;	Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
QY	1	CGCTGGAGGATCATCCAGCCGGCTCCCGGAAACGATCCGAGGCCAACCTTTCATAG	60	
DB	1390	CGCTGGAGGATCATCCAGCCGGCTCCCGGAAACGATCCGAGGCCAACCTTTCATAG	1331	
QY	61	AAGCGCGGTGGAAATCGAAATCTCGTATGGCAGGTGGGGCTGGCTTGGTTCGCTCATT	120	
DB	1330	AAGCGCGGTGGAAATCGAAATCTCGTATGGCAGGTGGGGCTGGCTTGGTTCGCTCATT	1271	
QY	121	TCGAACCCAGAGTCCCGCTCAGAAGAACTCGTCAAGAGGGCGATAGAGGCGATGCCCT	180	
DB	1270	TCGAACCCAGAGTCCCGCTCAGAAGAACTCGTCAAGAGGGCGATAGAGGCGATGCCCT	1211	
QY	181	GGGAATCGGAGCGCGGATACCGTAAAGCAAGAGGAGCGGTCAGCCCATTCGCGGCCAA	240	
DB	1210	GGGAATCGGAGCGCGGATACCGTAAAGCAAGAGGAGCGGTCAGCCCATTCGCGGCCAA	1151	
QY	241	GCTCTTCAGCAATATCACGGTAGCCAAACGCTATGTCCTGATAGCGGTCCGCCACACCCA	300	
DB	1150	GCTCTTCAGCAATATCACGGTAGCCAAACGCTATGTCCTGATAGCGGTCCGCCACACCCA	1091	
QY	301	GGCGGCCACAGTCGATGAATCCAGAAAAGCGCCATTTTCCACCATGATATTCGGCGAAGC	360	
DB	1090	GGCGGCCACAGTCGATGAATCCAGAAAAGCGCCATTTTCCACCATGATATTCGGCGAAGC	1031	
QY	361	AGGCATCGCCATGGGTCAACACAGAGATCCTCGCGCTCGGCGATCGCGCTTGAGCCTGG	420	
DB	1030	AGGCATCGCCATGGGTCAACACAGAGATCCTCGCGCTCGGCGATCGCGCTTGAGCCTGG	971	
QY	421	CGAACAGTTGGTGGCGGAGCCCTGATGCTCTTCGTCCAGATCATCTCTGATCGACAA	480	
DB	970	CGAACAGTTGGTGGCGGAGCCCTGATGCTCTTCGTCCAGATCATCTCTGATCGACAA	911	
QY	481	GACCGGTTCCATCCGAGTACGTGCTCGATCGATGCGATGCTTCGTCCAGATCATCTCTGATCGAATG	540	
DB	910	GACCGGTTCCATCCGAGTACGTGCTCGATCGATGCGATGCTTCGTCCAGATCATCTCTGATCGAATG	851	
QY	541	GGCAGGTAGCGGATCAAGCGTATGAGCGCGCGCATTTGCATCAGCCATGATGGATACCTT	600	
DB	850	GGCAGGTAGCGGATCAAGCGTATGAGCGCGCGCATTTGCATCAGCCATGATGGATACCTT	791	
QY	601	TCTCGCAGGAGCAAGGTGAGATGACAGGAGATCCTGCCCGGCATCTGCCCGAATAGCA	660	
DB	790	TCTCGCAGGAGCAAGGTGAGATGACAGGAGATCCTGCCCGGCATCTGCCCGAATAGCA	731	
QY	661	GCCAGTCTCTTCGGCTTCAGTGACAACTCGACACAGCTGCGCAAGGAACGCCCGCTCG	720	
DB	730	GCCAGTCTCTTCGGCTTCAGTGACAACTCGACACAGCTGCGCAAGGAACGCCCGCTCG	671	
QY	721	TGGCCAGCACAGTAGCCGCTGCTCTGCTGAGTTCATTCAGGGCACCGGACAGGT	780	
DB	670	TGGCCAGCACAGTAGCCGCTGCTCTGCTGAGTTCATTCAGGGCACCGGACAGGT	611	
QY	781	CGGTTCTTGACAAAAGAAACGGGCGCCCTTGGCTGACAGCCGGAAACACCGCGCATCAG	840	

RESULT 3

```

RESUL 3
US-09-027-287-50
; Sequence 50, Application US/09027287A
; Patent No. US20020064869A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ruben, Steven M.
; APPLICANT: Ullrich, Stephen
; TITLE OF INVENTION: Apoptosis Inducing Molecule II
; FILE REFERENCE: 1488.0650004
; CURRENT APPLICATION NUMBER: US/09/027,287A
; CURRENT FILING DATE: 1998-02-20
; EARLIER APPLICATION NUMBER: US 09/003,886
; EARLIER FILING DATE: 1998-01-07
; EARLIER APPLICATION NUMBER: US 08/822,953
; EARLIER FILING DATE: 1997-03-21
; EARLIER APPLICATION NUMBER: US 60/030,157
; EARLIER FILING DATE: 1996-10-31
; EARLIER APPLICATION NUMBER: US 60/013,923
; EARLIER FILING DATE: 1996-03-22
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50
; LENGTH: 3974
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: vector
US-09-027-287-50

```

Query Match	100.0%;	Score 1001;	DB 9;	Length 3974;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1001; Conservative	0;	Mismatches	0;	Indels
	0;	Gaps	0;	Gaps
QY	1	CGCTGAGGATCATCCAGCGGGGTC	CCGGAAACGATTCCG	GAAGCCACCTTTTCATAG 60
Db	2579	CGCTGAGGATCATCCAGCGGGGTC	CCGGAAACGATTCCG	GAAGCCACCTTTTCATAG 2638
QY	61	AAGCGCGCGGTGGAATCGAAATCT	CTGTGATGGCAGGTCGGGGT	TCGCTTGGTCGGTCATT 120
Db	2639	AAGCGCGCGGTGGAATCGAAATCT	CTGTGATGGCAGGTCGGGGT	TCGCTTGGTCGGTCATT 2698
QY	121	TGGAACCCAGAGTCCCGTTCAGAA	AACTCGTCAAGAAGGCGATAG	AGGCGATCGCT 180
Db	2699	TGGAACCCAGAGTCCCGTTCAGAA	AACTCGTCAAGAAGGCGATAG	AGGCGATCGCT 2758
QY	181	GGGAATCGGGAGCGGGATACCGT	TAAAGCACAGGAGCGGT	CAGCCCAATTCGCCGCCAA 240
Db	2759	GGGAATCGGGAGCGGGATACCGT	TAAAGCACAGGAGCGGT	CAGCCCAATTCGCCGCCAA 2818
QY	241	GCTCTTTAGCAATATCAGCGGT	AGCCAAACGCTATGCTGATAG	CGGTCGGGCACACCCA 300
Db	2819	GCTCTTTAGCAATATCAGCGGT	AGCCAAACGCTATGCTGATAG	CGGTCGGGCACACCCA 2878
QY	301	CGCGGCCACAGTCGATGAAATC	GAAAAAGCGGCCATTTTCCA	CCCAATGATATTCGGCAAGC 360
Db	2879	CGCGGCCACAGTCGATGAAATC	GAAAAAGCGGCCATTTTCCA	CCCAATGATATTCGGCAAGC 2938

Qy	361	AGG	CAT	GCC	CA	TGG	GT	CA	CAG	CAG	AAT	CCT	CG	CG	T	CGG	G	CAT	CG	CG	CCT	TG	AG	CCT	GG	420						
Db	2939	AGG	CAT	GCC	CA	TGG	GT	CAC	GAC	GAG	AT	CCT	CG	CG	T	CGG	G	CAT	CG	CG	CCT	TG	AG	CCT	GG	2998						
Qy	421	CGA	AC	AG	TTC	GG	CT	GG	CG	GAG	CC	CT	GAT	G	CT	CT	T	CG	T	CG	AG	AT	CAT	CCT	GAT	C	CAA	480				
Db	2999	CGA	AC	AG	TTC	GG	CT	GG	CG	GAG	CC	CT	GAT	G	CT	CT	T	CG	T	CG	AG	AT	CAT	CCT	GAT	C	CAA	3058				
Qy	481	GAC	CG	G	CT	T	CC	AT	CC	GAG	T	AG	CT	CG	T	CG	T	CG	A	T	GC	AT	TC	T	T	CG	CG	AT	G	540		
Db	3059	GAC	CG	G	CT	T	CC	AT	CC	GAG	T	AG	CT	CG	T	CG	T	CG	A	T	GC	AT	TC	T	T	CG	CG	AT	G	3118		
Qy	541	GGC	AG	T	AG	CG	GA	T	CA	AG	CG	T	AT	G	C	A	G	C	G	C	CA	T	T	G	CAT	CAG	CC	AT	G	600		
Db	3119	GGC	AG	T	AG	CG	GA	T	CA	AG	CG	T	AT	G	C	A	G	C	G	C	CA	T	T	G	CAT	CAG	CC	AT	G	3178		
Qy	601	TCT	CG	G	CAG	GAC	AA	GT	GAG	AT	GAC	GAG	AT	CCT	CG	CG	G	CA	CT	T	CG	CG	CA	CT	T	CG	CG	CA	AT	G	660	
Db	3179	TCT	CG	G	CAG	GAC	AA	GT	GAG	AT	GAC	GAG	AT	CCT	CG	CG	G	CA	CT	T	CG	CG	CA	CT	T	CG	CG	CA	AT	G	3238	
Qy	661	GCC	AG	T	CC	CT	T	CG	CT	T	CAG	T	GAC	AA	CG	T	CG	A	C	A	G	C	A	G	C	A	G	C	A	G	720	
Db	3239	GCC	AG	T	CC	CT	T	CG	CT	T	CAG	T	GAC	AA	CG	T	CG	A	C	A	G	C	A	G	C	A	G	C	A	G	3298	
Qy	721	TGG	CC	AG	CC	AC	CA	T	AG	CG	CG	T	GC	CT	CG	CG	A	CT	T	CA	T	T	CAG	GC	CG	CG	CG	CG	CG	CG	780	
Db	3299	TGG	CC	AG	CC	AC	CA	T	AG	CG	CG	T	GC	CT	CG	CG	A	CT	T	CA	T	T	CAG	GC	CG	CG	CG	CG	CG	CG	3358	
Qy	781	CGG	CT	T	G	CA	AA	AA	GA	AA	CC	GG	CG	CC	CT	T	GC	CT	G	A	C	A	G	C	CG	GA	AA	CA	CG	GG	CG	840
Db	3359	CGG	CT	T	G	CA	AA	AA	GA	AA	CC	GG	CG	CC	CT	T	GC	CT	G	A	C	A	G	C	CG	GA	AA	CA	CG	GG	CG	3418
Qy	841	AG	CAG	CC	GA	T	T	G	T	G	T	G	CC	CA	T	AG	CC	GA	T	AG	CC	T	CT	C	CA	CC	AA	CG	CG	CG	900	
Db	3419	AG	CAG	CC	GA	T	T	G	T	G	T	G	CC	CA	T	AG	CC	GA	T	AG	CC	T	CT	C	CA	CC	AA	CG	CG	CG	3478	
Qy	901	GAG	AA	C	T	GG	T	GC	AA	T	CT	T	GT	T	CA	T	AT	C	AT	G	A	A	C	GA	T	CT	C	A	T	CT	T	960
Db	3479	GAG	AA	C	T	GG	T	GC	AA	T	CT	T	GT	T	CA	T	AT	C	AT	G	A	A	C	GA	T	CT	C	A	T	CT	T	3538
Qy	961	GAT	CAG	AT	TT	GA	T	CC	CT	T	GC	CG	CA	T	CAG	AT	CT	T	T	GG	CG	C									1001	
Db	3539	GAT	CAG	AT	TT	GA	T	CC	CT	T	GC	CG	CA	T	CAG	AT	CT	T														

RESULT 4

```

US-09-252-656B-50
; Sequence 50, Application US/09252656B
; Patent No. US20020081647A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ruben, Steven M.
; APPLICANT: Zhang, Jun
; APPLICANT: Ullrich, Stephen
; APPLICANT: Zhai, Yifan
; TITLE OF INVENTION: Apoptosis Inducing Molecule II and Methods of Use
; FILE REFERENCE: 1488.0650006
; CURRENT APPLICATION NUMBER: US/09/252.656B
; CURRENT FILING DATE: 1993-02-19
; PRIOR APPLICATION NUMBER: US 60/075,409
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 09/027,287
; PRIOR FILING DATE: 1998-02-20
; PRIOR APPLICATION NUMBER: US 09/003,886
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: US 08/822,953
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/013,923
; PRIOR FILING DATE: 1996-03-22
; PRIOR APPLICATION NUMBER: US 60/030,157
; PRIOR FILING DATE: 1996-10-31

```

QY 901 GAGACCTCGGTGCAATCCATCTTGTTCATCATCGGAAACGATCCCTCTCTCTT 960
Db 3479 GAGACCTCGGTGCAATCCATCTTGTTCATCATCGGAAACGATCCCTCTCTCTT 3538
QY 961 GATCAGATCTTGATCCCTCGGCAATCAGATCCCTTGGCGGC 1001
Db 3539 GATCAGATCTTGATCCCTCGGCAATCAGATCCCTTGGCGGC 3579

RESULT 5
US-09-972-912-8
; Sequence 8, Application US/09972912
; Patent No. US20020110867A1
; GENERAL INFORMATION:
; APPLICANT: SOPPET, DANIEL R.
; RUBEN, STEVEN M.
; TITLE OF INVENTION: CARDIAC AND PANCREATIC PROTEIN AND GENE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: US
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/972,912
; FILING DATE: 10-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/049,022
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0620001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cdna
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-972-912-8

Query Match 100.0%; Score 1001; DB 10; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGGAGGATCATCCAGCGCGGTCCCGGAAACGATTCGGAAGCCCAACCTTTCATAG 60
Db 2579 CGTGGAGGATCATCCAGCGCGGTCCCGGAAACGATTCGGAAGCCCAACCTTTCATAG 2638
QY 61 AAGCGCGGTGGAATCGAATCTCGTATGCGAGGTGGCGGTGCGTTCGTCGTCATT 120
Db 2639 AAGCGCGGTGGAATCGAATCTCGTATGCGAGGTGGCGGTGCGTTCGTCGTCATT 2698
QY 121 TCGAACCACAGAGTCCCGCTCAGAAGACTCGTCAAGAGGCGATAGAGGCGGATCGCT 180
Db 2699 TCGAACCACAGAGTCCCGCTCAGAAGACTCGTCAAGAGGCGATAGAGGCGGATCGCT 2759
QY 181 GCGAATCGGAGGCGGATACCGTAAAGCAGAGGAGCGGTACGCCCAATCCGCCCAAA 240

Db 2759 GCGAATCGGAGGCGGATACCGTAAAGCAGAGGAGCGGTACGCCCAATCCGCCCAAA 2818
QY 241 GCTCTTCAGCAATATCACGGGTAGCAACGCTATGTCTGTATAGCGGTCCGCCACACCCA 300
Db 2819 GCTCTTCAGCAATATCACGGGTAGCAACGCTATGTCTGTATAGCGGTCCGCCACACCCA 2878
QY 301 GCGGCGCACAGTCGATGAATCCAGAAAAGCGGCCATTTTCCACCATGATATTCGCAAGC 360
Db 2879 GCGGCGCACAGTCGATGAATCCAGAAAAGCGGCCATTTTCCACCATGATATTCGCAAGC 2938
QY 361 AGGCATCGCATGGGTACGACGAGATCCTCGCCGTGCGGATCGCGCTTGGCCTTGG 420
Db 2939 AGGCATCGCATGGGTACGACGAGATCCTCGCCGTGCGGATCGCGCTTGGCCTTGG 2998
QY 421 CGAACAGTTCCGCTCGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTTGATGACAA 480
Db 2999 CGAACAGTTCCGCTCGCGGAGCCCTGATGCTCTTCTGTCAGATCATCTTGATGACAA 3058
QY 481 GACCGGCTTCCATCCGAGTACGTGCTCGCTCGATCGAGTGTTCGCTTGGTGGTCAATG 540
Db 3059 GACCGGCTTCCATCCGAGTACGTGCTCGCTCGATCGAGTGTTCGCTTGGTGGTCAATG 3118
QY 541 GCGAGTAGCCGATCAAGCGTATGACGCGCGCGCAATGTCATCAGCCATGATGATCTT 600
Db 3119 GCGAGTAGCCGATCAAGCGTATGACGCGCGCGCAATGTCATCAGCCATGATGATCTT 3178
QY 601 TCTCGCAGGAGCAAGGTGAGATGACAGGAGATCTGCCCCGGCACTTCGCCCAATAGCA 660
Db 3179 TCTCGCAGGAGCAAGGTGAGATGACAGGAGATCTGCCCCGGCACTTCGCCCAATAGCA 3238
QY 661 GCGAGTCCCTTCCCGCTTCAAGCAAGCTGAGCAGACAGTCCGCAAGGAGCGCCGCG 720
Db 3239 GCGAGTCCCTTCCCGCTTCAAGCAAGCTGAGCAGACAGTCCGCAAGGAGCGCCGCG 3298
QY 721 TGGCCAGCCACGATAGCGCGCTGCTGCTCTGCTGCTTCAATCAGGCGACCGGACAGT 780
Db 3299 TGGCCAGCCACGATAGCGCGCTGCTGCTCTGCTGCTTCAATCAGGCGACCGGACAGT 3358
QY 781 CGGTCTTGACAAAAGAACCGGGCGCCCTGCGGTGACAGCGGACACGCGGCGCATCAG 840
Db 3359 CGGTCTTGACAAAAGAACCGGGCGCCCTGCGGTGACAGCGGACACGCGGCGCATCAG 3418
QY 841 AGCAGCCGATTTGCTGTTGTGTCAGTCATAGCCGAATAGCTCTCCACCCCAAGCGCG 900
Db 3419 AGCAGCCGATTTGCTGTTGTGTCAGTCATAGCCGAATAGCTCTCCACCCCAAGCGCG 3478
QY 901 GAGAACCTGCGTCAATCCATCTTGTTCATCATGCGAAGCATCTCTCATCTCTCTT 960
Db 3479 GAGAACCTGCGTCAATCCATCTTGTTCATCATGCGAAGCATCTCTCATCTCTCTT 3538
QY 961 GATCAGATCTTGATCCCTCGGCAATCAGATCCCTTGGCGGC 1001
Db 3539 GATCAGATCTTGATCCCTCGGCAATCAGATCCCTTGGCGGC 3579

RESULT 6

US-09-902-684-14
; Sequence 14, Application US/09902684
; Patent No. US20020127640A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: PANCREAS-DERIVED PLASMINOGEN ACTIVATOR INHIBITOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/902,684
 FILING DATE: 12-Jul-2001
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/026,408
 FILING DATE: <Unknown>
 APPLICATION NUMBER: US 60/024,056
 FILING DATE: 16-AUG-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: STEFFE, ERIC K.
 REGISTRATION NUMBER: 36,688
 REFERENCE/DOCKET NUMBER: 1488.03000002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-371-2600
 TELEFAX: 202-371-2540
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3974 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: both
 TOPOLOGY: both
 MOLECULE TYPE: cDNA
 SEQUENCE DESCRIPTION: SEQ ID NO: 14:
 US-09-902-684-14

Query Match 100.0%; Score 1001; DB 10; Length 3974;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGCTGAGGATCATCCAGCGCGCTCCCGGAAAGCATTCGAAAGCCCAACCTTTCATAG 60
 DB 2579 CGCTGAGGATCATCCAGCGCGCTCCCGGAAAGCATTCGAAAGCCCAACCTTTCATAG 2638
 QY 61 AAGGCGCGGTGGATCGAAATCTGTGATGCGAGTGGCGTGGCGTGGCGTGGCGTGGCGT 120
 DB 2639 AAGGCGCGGTGGATCGAAATCTGTGATGCGAGTGGCGTGGCGTGGCGTGGCGTGGCGT 2698
 QY 121 TCGAAGCCCGAGAGTCCCGTTCAGAAAGTCTCGTCAAGAGCGGATAGAGCGGATGGCT 180
 DB 2699 TCGAAGCCCGAGAGTCCCGTTCAGAAAGTCTCGTCAAGAGCGGATAGAGCGGATGGCT 2758
 QY 181 GCGAATCGGAGCGGCGATACCGTAAAGCAAGAGCGGTCAGCCCATTCGCGGCAA 240
 DB 2759 GCGAATCGGAGCGGCGATACCGTAAAGCAAGAGCGGTCAGCCCATTCGCGGCAA 2818
 QY 241 GCTCTTCAGCAATATACCGGTAGCAACGCTATGCTGATAGCGGTCCGCGCACCCA 300
 DB 2819 GCTCTTCAGCAATATACCGGTAGCAACGCTATGCTGATAGCGGTCCGCGCACCCA 2878
 QY 301 GCGGCGCAGTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATATTCGGCAAGC 360
 DB 2879 GCGGCGCAGTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATATTCGGCAAGC 2938
 QY 361 AGGCATCGGCATGGGTACGACGATCTCGCGTGGCGATGGCGGCTTGGAGCTGG 420
 DB 2939 AGGCATCGGCATGGGTACGACGATCTCGCGTGGCGATGGCGGCTTGGAGCTGG 2998
 QY 421 CGAAAGTTCGGCTGGCGGAGCCCTTGATGCTCTTCTGTCAGATCATCTCTGATCGACAA 480
 DB 2999 CGAAAGTTCGGCTGGCGGAGCCCTTGATGCTCTTCTGTCAGATCATCTCTGATCGACAA 3058
 QY 481 GACCGGCTTCATCCGAGTGGTCTGCTCGATGCGATGTTGCTGGTGGTGGTGGTGGT 540
 DB 3059 GACCGGCTTCATCCGAGTGGTCTGCTCGATGCGATGTTGCTGGTGGTGGTGGTGGT 3118
 QY 541 GCGAGTACCGGATCAAGCGTATGACCGCGCGCATTCGATCAGCCATGATGATGATGAT 600
 DB 3119 GCGAGTACCGGATCAAGCGTATGACCGCGCGCATTCGATCAGCCATGATGATGATGAT 3178

QY 601 TCTCGGAGGAGCAAGCTGAGATGACAGGAGATCTCTCCCGGACATTCGCCCAATAGCA 660
 DB 3179 TCTCGGAGGAGCAAGCTGAGATGACAGGAGATCTCTCCCGGACATTCGCCCAATAGCA 3238
 QY 661 GCGAGTCCCTTCCCGCTTCAGTGACAGCTGAGACAGCTGCGCAAGGAAAGCGCCCTCG 720
 DB 3239 GCGAGTCCCTTCCCGCTTCAGTGACAGCTGAGACAGCTGCGCAAGGAAAGCGCCCTCG 3298
 QY 721 TGGCCAGGACGATAGCGCGCTCGCTTCGCTGAGTTCATTCAGGCGACCGGACAGGT 780
 DB 3299 TGGCCAGGACGATAGCGCGCTCGCTTCGCTGAGTTCATTCAGGCGACCGGACAGGT 3358
 QY 781 CGCTCTTGACAAAAAGAACCGGCGGCCCTTGGCGTGACAGCCGAAACACCGCGGATCAG 840
 DB 3359 CGCTCTTGACAAAAAGAACCGGCGGCCCTTGGCGTGACAGCCGAAACACCGCGGATCAG 3418
 QY 841 AGCAGCGGATGCTGTGTGTCGCCAGTCATAGCCGATAGCCCTCTCACCAAGCGGCG 900
 DB 3419 AGCAGCGGATGCTGTGTGTCGCCAGTCATAGCCGATAGCCCTCTCACCAAGCGGCG 3478
 QY 901 GAGAACCTGCTGCAATCCATCTTGTTCATATGAGGAAACGATCCTCATCTGCTCTT 960
 DB 3479 GAGAACCTGCTGCAATCCATCTTGTTCATATGAGGAAACGATCCTCATCTGCTCTT 3538
 QY 961 GATCAGATCTTGCCTCCGCGCATCAGATCCTTGGCGGC 1001
 DB 3539 GATCAGATCTTGCCTCCGCGCATCAGATCCTTGGCGGC 3579

RESULT 7
 US-09-935-726-16
 ; Sequence 16, Application US/09935726
 ; Publication No. US20030008357A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hu, Jin-Shan
 ; APPLICANT: Craig, Rosen
 ; APPLICANT: Cao, Liang
 ; TITLE OF INVENTION: Vascular Endothelial Growth Factor-2
 ; FILE REFERENCE: PF112P3D1C1
 ; CURRENT APPLICATION NUMBER: US/09/935,726
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: 09/438,538
 ; PRIOR FILING DATE: 1999-11-12
 ; NUMBER OF SEQ ID NOS: 35
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 16
 ; LENGTH: 3974
 ; TYPE: DNA
 ; ORGANISM: Escherichia coli
 US-09-935-726-16

Query Match 100.0%; Score 1001; DB 11; Length 3974;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGCTGAGGATCATCCAGCGCGCTCCCGGAAAGCATTCGAAAGCCCAACCTTTCATAG 60
 DB 2579 CGCTGAGGATCATCCAGCGCGCTCCCGGAAAGCATTCGAAAGCCCAACCTTTCATAG 2638
 QY 61 AAGGCGCGGTGGATCGAAATCTCGTATGCGAGTGGCGTGGCGTGGCGTGGCGTGGCGT 120
 DB 2639 AAGGCGCGGTGGATCGAAATCTCGTATGCGAGTGGCGTGGCGTGGCGTGGCGTGGCGT 2698
 QY 121 TCGAAGCCCGAGAGTCCCGTTCAGAAAGTCTCGTCAAGAGCGGATAGAGCGGATGGCT 180
 DB 2699 TCGAAGCCCGAGAGTCCCGTTCAGAAAGTCTCGTCAAGAGCGGATAGAGCGGATGGCT 2758
 QY 181 GCGAATCGGAGCGGCGATACCGTAAAGCAAGAGCGGTCAGCCCATTCGCGGCAA 240
 DB 2759 GCGAATCGGAGCGGCGATACCGTAAAGCAAGAGCGGTCAGCCCATTCGCGGCAA 2818
 QY 241 GCTCTTCAGCAATATACCGGTAGCAACGCTATGCTGATAGCGGTCCGCGCACCCA 300

Db 2819 GCTCTTACGAAATATACGGGTAGCAAGCGTATGCTCTGATAGCGGTCCGCCACACCA 2878
QY 301 GCGGCGCAGTCTGATGATCCAGAAAGCGGCCATTTTCCACCATGATATTCGGCAAGC 360
Db 2879 GCGGCGCAGTCTGATGATCCAGAAAGCGGCCATTTTCCACCATGATATTCGGCAAGC 2938
QY 361 AGGCATCGCCATGCGGTACGACGAGATCCTGCGGTGGGATGCGGCGCTTGAGCGCTGG 420
Db 2939 AGGCATCGCCATGCGGTACGACGAGATCCTGCGGTGGGATGCGGCGCTTGAGCGCTGG 2998
QY 421 CGAACAGTTGCGTGGCGGAGCGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 480
Db 2999 CGAACAGTTGCGTGGCGGAGCGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 3058
QY 481 GACCGGCTTCCATCGGATGATGCTGCTCGATGCGGATGCTTGGTGGTGGTGGTGGTGG 540
Db 3059 GACCGGCTTCCATCGGATGATGCTGCTCGATGCGGATGCTTGGTGGTGGTGGTGGTGG 3118
QY 541 GCGAGGTAGCGGATCAAGCGTATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATG 600
Db 3119 GCGAGGTAGCGGATCAAGCGTATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATG 3178
QY 601 TCTGCGGAGGACGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Db 3179 TCTGCGGAGGACGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3238
QY 661 GCGAGTCTCTTCCGCGTTCAGTGACAACTGTCAGACAGTTCGCAAGGAGCGCGGTGG 720
Db 3239 GCGAGTCTCTTCCGCGTTCAGTGACAACTGTCAGACAGTTCGCAAGGAGCGCGGTGG 3298
QY 721 TGCGGAGGACGATAGCGGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 3299 TGCGGAGGACGATAGCGGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3358
QY 781 CGGTCTTGACAAAAGAACCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 3359 CGGTCTTGACAAAAGAACCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3418
QY 841 AGGAGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 3419 AGGAGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3478
QY 901 GAGAACCTGCTGCAATCCATCTGTTCAATCATGCGAAGCATGCTTCCACCCAGCGCGCG 960
Db 3479 GAGAACCTGCTGCAATCCATCTGTTTCAATCATGCGAAGCATGCTTCCACCCAGCGCG 3538
QY 961 GATCAGATCTTGATCCCTCGCCATCAGATCCTTGGCGGC 1001
Db 3539 GATCAGATCTTGATCCCTCGCCATCAGATCCTTGGCGGC 3579

RESULT 8
US-09-105-470-25
; Sequence 25, Application US/09105470
; Publication No. US2003002275A1
; GENERAL INFORMATION:
; APPLICANT: Carter, Kenneth C.
; APPLICANT: He, Wei-Wu
; TITLE OF INVENTION: Human NK-3 Related Prostate Specific
; TITLE OF INVENTION: Gene-1
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,470
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/051,080
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0790001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3974 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; US-09-105-470-25

Query Match 100.0%; Score 1001; DB 11; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGTGGAGGATCATCCAGCGCGGTCCCGGAAACGATTCGGAAGCCCAACCTTTTCATAG 60
Db 2579 CCGTGGAGGATCATCCAGCGCGGTCCCGGAAACGATTCGGAAGCCCAACCTTTTCATAG 2638
QY 61 AAGGCGGCGGTGGAATCTGTCATGCGAGGTTGGGCGTCTGCTTGGTGGTCAAT 120
Db 2639 AAGGCGGCGGTGGAATCTGTCATGCGAGGTTGGGCGTCTGCTTGGTGGTCAAT 2698
QY 121 TCGAACCCCGAGTCCCCTCAGAGAACTCGTCAAGAGGCGATAGAGGCGGATCGCCT 180
Db 2699 TCGAACCCCGAGTCCCCTCAGAGAACTCGTCAAGAGGCGATAGAGGCGGATCGCCT 2758
QY 181 GCGAATCGGAGCGCGGATAGCGTAAGACAGAGAGGCGGTGAGCCCATTCGCGGCCAA 240
Db 2759 GCGAATCGGAGCGCGGATAGCGTAAGACAGAGAGGCGGTGAGCCCATTCGCGGCCAA 2818
QY 241 GCTCTTACGAAATATCAGCGGTAGCCACGCTATGCTCTGATAGCGGTCCGCCACACCCA 300
Db 2819 GCTCTTACGAAATATCAGCGGTAGCCACGCTATGCTCTGATAGCGGTCCGCCACACCCA 2878
QY 301 GCGGCCACAGTCGATGATCCAGAAAGCGGCCATTTTCCACCATGATATTCGGCAAGC 360
Db 2879 GCGGCCACAGTCGATGATCCAGAAAGCGGCCATTTTCCACCATGATATTCGGCAAGC 2938
QY 361 AGGCATCGCCATGCGGTACGACGAGATCCTGCGGTGGGATGCGGCGCTTGAGCGCTGG 420
Db 2939 AGGCATCGCCATGCGGTACGACGAGATCCTGCGGTGGGATGCGGCGCTTGAGCGCTGG 2998
QY 421 CGAACAGTTGCGTGGCGGAGCGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 480
Db 2999 CGAACAGTTGCGTGGCGGAGCGCCCTGATGCTCTTCTGTCAGATCATCTGATCGACAA 3058
QY 481 GACCGGCTTCCATCGGATGATGCTGCTCGATGCGGATGCTTGGTGGTGGTGGTGGTGG 540
Db 3059 GACCGGCTTCCATCGGATGATGCTGCTCGATGCGGATGCTTGGTGGTGGTGGTGGTGG 3118
QY 541 GCGAGGTAGCGGATCAAGCGTATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATG 600
Db 3119 GCGAGGTAGCGGATCAAGCGTATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATG 3178
QY 601 TCTGCGGAGGACGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Db 3179 TCTGCGGAGGACGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3238
QY 661 GCGAGTCTCTTCCGCGTTCAGTGACAACTGTCAGACAGTTCGCAAGGAGCGCGGTGG 720

Qy 61 AAGCGCGGTGGGAATCGAATCTCTGATGAGGAGTGGCGTCTGGTCTGGTCTGCTT 120
Db 2639 AAGCGCGGTGGGAATCGAATCTCTGATGAGGAGTGGCGTCTGGTCTGGTCTGCTT 2698
Qy 121 TCGAACCCCGAGTCCCGCTCAGAGNACTCTGTCAGAGAGGCGATAGAGGCGATCCGCT 180
Db 2699 TCGAACCCCGAGTCCCGCTCAGAGNACTCTGTCAGAGAGGCGATAGAGGCGATCCGCT 2758
Qy 181 GCGAATCGGAGCGCGATACCGTAAAGCAGGAGGCGGTGAGCCCATTCGCGCCAA 240
Db 2759 GCGAATCGGAGCGCGATACCGTAAAGCAGGAGGCGGTGAGCCCATTCGCGCCAA 2818
Qy 241 GCTCTTCAGCAATATCAGCGGTAGCCAGCTATGCTCTGATAGCGTCCGCGCACCA 300
Db 2819 GCTCTTCAGCAATATCAGCGGTAGCCAGCTATGCTCTGATAGCGTCCGCGCACCA 2878
Qy 301 GCGCGCCACAGTCGATGAATCCAGAAAGCGGCGCATTTTCCACCATGATATTCGCGAAGC 360
Db 2879 GCGCGCCACAGTCGATGAATCCAGAAAGCGGCGCATTTTCCACCATGATATTCGCGAAGC 2938
Qy 361 AGGATCGCGATGGTCAAGCAGGATCTCTGCGCGTGGCGATGCGCGCTTGGCGCTGG 420
Db 2939 AGGATCGCGATGGTCAAGCAGGATCTCTGCGCGTGGCGATGCGCGCTTGGCGCTGG 2998
Qy 421 CGAACAGTTCGGCTGGCGCGAGCCCTGATCTCTCTGTCAGATCATCTCTGATCGACAA 480
Db 2999 CGAACAGTTCGGCTGGCGCGAGCCCTGATCTCTCTGTCAGATCATCTCTGATCGACAA 3058
Qy 481 GACCGGCTTCATCCGAGTACGCTCTGCTCGATGCGATGTTTGGCTTGGTGGTGAATG 540
Db 3059 GACCGGCTTCATCCGAGTACGCTCTGCTCGATGCGATGTTTGGCTTGGTGGTGAATG 3118
Qy 541 GCGAGTGGCGGATCAAGCGGTATGACAGCGCGCGATTCATCAGCCATGATGATATCT 600
Db 3119 GCGAGTGGCGGATCAAGCGGTATGACAGCGCGCGATTCATCAGCCATGATGATATCT 3178
Qy 601 TCTGCGAGGAGGAGTGAATGACAGAGATCTCTGCTGATGCGATGTTTGGCTTGGTGGTGAATG 660
Db 3179 TCTGCGAGGAGGAGTGAATGACAGAGATCTCTGCTGATGCGATGTTTGGCTTGGTGGTGAATG 3238
Qy 661 GCGAGTGGCGGATCAAGCGGTATGACAGCGCGCGATTCATCAGCCATGATGATATCT 720
Db 3239 GCGAGTGGCGGATCAAGCGGTATGACAGCGCGCGATTCATCAGCCATGATGATATCT 3298
Qy 721 TGGCGAGGAGGAGTGAATGACAGAGATCTCTGCTGATGCGATGTTTGGCTTGGTGGTGAATG 780
Db 3299 TGGCGAGGAGGAGTGAATGACAGAGATCTCTGCTGATGCGATGTTTGGCTTGGTGGTGAATG 3358
Qy 781 GCGAGTGGCGGATCAAGCGGTATGACAGCGCGCGATTCATCAGCCATGATGATATCT 840
Db 3359 GCGAGTGGCGGATCAAGCGGTATGACAGCGCGCGATTCATCAGCCATGATGATATCT 3418
Qy 841 AGCAGCGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 3419 AGCAGCGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3478
Qy 901 GAGAACCTGCTGCAATCCATCTTGTTCATATGCGAAACGATCTCTGCTGCTGCTGCT 960
Db 3479 GAGAACCTGCTGCAATCCATCTTGTTCATATGCGAAACGATCTCTGCTGCTGCTGCT 3538
Qy 961 GATCAGATCTTGTATCCCGTGGCGCATCAGATCTCTGCGGC 1001
Db 3539 GATCAGATCTTGTATCCCGTGGCGCATCAGATCTCTGCGGC 3579

RESULT 12

US-10-375-680-50
; Sequence 50, Application US/10375680
; Publication No. US20040009147A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ruben, Steven M

; APPLICANT: Ullrich, Stephen
; APPLICANT: Zhai, Yifan
; TITLE OF INVENTION: Apoptosis Inducing Molecule II and Methods of Use
; FILE REFERENCE: 1488 065000E
; CURRENT APPLICATION NUMBER: US/10375,680
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,234
; PRIOR FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 50
; LENGTH: 3974
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Vector
US-10-375-680-50

Query Match 100.0%; Score 1001; DB 12; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGTGGAGGATCATCCAGCGCGCTCCCGGAAAACGATCCGAGCCCAACCTTTCATAG 60
Db 2579 CGTGGAGGATCATCCAGCGCGCTCCCGGAAAACGATCCGAGCCCAACCTTTCATAG 2638
Qy 61 AAGCGCGGTGGGAATCGAATCTCTGATGAGGAGTGGCGTCTGGTCTGGTCTGCTT 120
Db 2639 AAGCGCGGTGGGAATCGAATCTCTGATGAGGAGTGGCGTCTGGTCTGGTCTGCTT 2698
Qy 121 TCGAACCCCGAGTCCCGCTCAGAGNACTCTGTCAGAGAGGCGATAGAGGCGATCCGCT 180
Db 2699 TCGAACCCCGAGTCCCGCTCAGAGNACTCTGTCAGAGAGGCGATAGAGGCGATCCGCT 2758
Qy 181 GCGAATCGGAGCGCGATACCGTAAAGCAGGAGGCGGTGAGCCCATTCGCGCCAA 240
Db 2759 GCGAATCGGAGCGCGATACCGTAAAGCAGGAGGCGGTGAGCCCATTCGCGCCAA 2818
Qy 241 GCTCTTCAGCAATATCAGCGGTAGCCAGCTATGCTCTGATAGCGTCCGCGCACCA 300
Db 2819 GCTCTTCAGCAATATCAGCGGTAGCCAGCTATGCTCTGATAGCGTCCGCGCACCA 2878
Qy 301 GCGCGCCACAGTCGATGAATCCAGAAAGCGGCGCATTTTCCACCATGATATTCGCGAAGC 360
Db 2879 GCGCGCCACAGTCGATGAATCCAGAAAGCGGCGCATTTTCCACCATGATATTCGCGAAGC 2938
Qy 361 AGGATCGCGATGGTCAAGCAGGATCTCTGCTGATGCGATGTTTGGCTTGGTGGTGAATG 420
Db 2939 AGGATCGCGATGGTCAAGCAGGATCTCTGCTGATGCGATGTTTGGCTTGGTGGTGAATG 2998
Qy 421 CGAACAGTTCGGCTGGCGCGAGCCCTGATCTCTCTGTCAGATCATCTCTGATCGACAA 480
Db 2999 CGAACAGTTCGGCTGGCGCGAGCCCTGATCTCTCTGTCAGATCATCTCTGATCGACAA 3058
Qy 481 GACCGGCTTCATCCGAGTACGCTCTGCTCGATGCGATGTTTGGCTTGGTGGTGAATG 540
Db 3059 GACCGGCTTCATCCGAGTACGCTCTGCTCGATGCGATGTTTGGCTTGGTGGTGAATG 3118
Qy 541 GCGAGTGGCGGATCAAGCGGTATGACAGCGCGCGATTCATCAGCCATGATGATATCT 600
Db 3119 GCGAGTGGCGGATCAAGCGGTATGACAGCGCGCGATTCATCAGCCATGATGATATCT 3178
Qy 601 TCTGCGAGGAGGAGTGAATGACAGAGATCTCTGCTGATGCGATGTTTGGCTTGGTGGTGAATG 660
Db 3179 TCTGCGAGGAGGAGTGAATGACAGAGATCTCTGCTGATGCGATGTTTGGCTTGGTGGTGAATG 3238
Qy 661 GCGAGTGGCGGATCAAGCGGTATGACAGCGCGCGATTCATCAGCCATGATGATATCT 720
Db 3239 GCGAGTGGCGGATCAAGCGGTATGACAGCGCGCGATTCATCAGCCATGATGATATCT 3298
Qy 721 TGGCGAGGAGGAGTGAATGACAGAGATCTCTGCTGATGCGATGTTTGGCTTGGTGGTGAATG 780
Db 3299 TGGCGAGGAGGAGTGAATGACAGAGATCTCTGCTGATGCGATGTTTGGCTTGGTGGTGAATG 3358

781 CGGTCTTGACAAAAGAACCGCGCCCTGGCTGACAGCCGGAACACCGCGGCATCAG 840
3359 CGGTCTTGACAAAAGAACCGCGCCCTGGCTGACAGCCGGAACACCGCGGCATCAG 3418
841 AGCAGCCGATGTCTGTTGCGCCAGTATAGCCGAATAGCTCTTCCACCCCAAGCGGCG 900
3419 AGCAGCCGATGTCTGTTGCGCCAGTATAGCCGAATAGCTCTTCCACCCCAAGCGGCG 3478
901 GAGAACCTGCGTGAATCCATCTTGTTCAATCATCGGAACGATCTCTCTCTCTT 960
3479 GAGAACCTGCGTGAATCCATCTTGTTCAATCATCGGAACGATCTCTCTCTCTT 3538
961 GATCAGATCTTGAATCCCTGCGCCATCAGATCCCTTGGCGGC 1001
3539 GATCAGATCTTGAATCCCTGCGCCATCAGATCCCTTGGCGGC 3579

RESULT 13
US-10-614-275-25
; Sequence 25, Application US/10614275
; Publication No. US20040014668A1
; GENERAL INFORMATION:
; APPLICANT: Carter, Kenneth C.
; He, Wei-Wu
; TITLE OF INVENTION: Human NK-3 Related Prostate Specific
; Gene-1
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/614,275
; FILING DATE: 08-Jul-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,470B
; FILING DATE: 26-Jun-1998
; APPLICATION NUMBER: US 60/051,080
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0790001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3974 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-10-614-275-25

Query Match 100.0%; Score 1001; DB 12; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 CGGTGAGGATCATCCAGCCGCTCCCGGAAACGATTCGGAAGCCCAACCTTTCATAG 60
2579 CGGTGAGGATCATCCAGCCGCTCCCGGAAACGATTCGGAAGCCCAACCTTTCATAG 2638

61 AAGCGCGCGTGGGAATCGAATCTCTGTGATGGCAGGTTGGCGCTCGCTTGGTGGTCACTT 120
2639 AAGCGCGCGTGGGAATCGAATCTCTGTGATGGCAGGTTGGCGCTCGCTTGGTGGTCACTT 2698
121 TCGAACCCCAAGATCCCGCTCAGAAAGAACTCGTCAAGAAAGCGGATAGAAAGCGGATGGCT 180
2699 TCGAACCCCAAGATCCCGCTCAGAAAGAACTCGTCAAGAAAGCGGATAGAAAGCGGATGGCT 2758
181 GCGAATCGGAGCGCGGATACCGTAAAGACGAGAGAAAGCGGTAGCGCCCATTCGCGCGCAA 240
2759 GCGAATCGGAGCGCGGATACCGTAAAGACGAGAGAAAGCGGTAGCGCCCATTCGCGCGCAA 2818
241 GCTCTTCAGCAATATACCGGTAGCCAAACGCTATGTCTCTGATAGCGGTTCGCCACACCCA 300
2819 GCTCTTCAGCAATATACCGGTAGCCAAACGCTATGTCTCTGATAGCGGTTCGCCACACCCA 2878
301 GCGCGCCACAGTTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATATTCGGCAAGC 360
2879 GCGCGCCACAGTTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATATTCGGCAAGC 2938
361 AGGCATCGCCATGGGTACAGCAGAGATCTCGCGCTCGGCGATGCGCGCTTGGTGGTGG 420
2939 AGGCATCGCCATGGGTACAGCAGAGATCTCGCGCTCGGCGATGCGCGCTTGGTGGTGG 2998
421 CGAACAGTTTCGCTGGGCGAGCCCTGATGCTCTTCGTCCAGATCATCTGATCGACAA 480
2999 CGAACAGTTTCGCTGGGCGAGCCCTGATGCTCTTCGTCCAGATCATCTGATCGACAA 3058
481 GACCGCTTCATCCAGTACGTCTCGCTCGATGCGATGTTTGGTGGTGGTGGTGGATG 540
3059 GACCGCTTCATCCAGTACGTCTCGCTCGATGCGATGTTTGGTGGTGGTGGTGGATG 3118
541 GCGAGTAGCCGATCAAGCGTATGACGCGCGCATTTGATCAGCCATGATGATGATCTT 600
3119 GCGAGTAGCCGATCAAGCGTATGACGCGCGCATTTGATCAGCCATGATGATGATCTT 3178
601 TCTCGCAGGAGCAAGGTGAGATCAGAGGATCTCTCGCGCGGCACTTCGCCCAATAGCA 660
3179 TCTCGCAGGAGCAAGGTGAGATCAGAGGATCTCTCGCGCGGCACTTCGCCCAATAGCA 3238
661 GCGAGTCCCTTCCCGCTTCAGTGACACAGTGGCAGCAGTGGCAGGAAACCGCGCTCG 720
3239 GCGAGTCCCTTCCCGCTTCAGTGACACAGTGGCAGCAGTGGCAGGAAACCGCGCTCG 3298
721 TGGCAGCCACGATAGCGCGCTCGCTCGCTGAGTTCAATCAGGCGCACCGGACAGGT 780
3299 TGGCAGCCACGATAGCGCGCTCGCTCGCTGAGTTCAATCAGGCGCACCGGACAGGT 3358
781 CGGTCTTGACAAAAGAACCGCGCCCTGGCTGACAGCCGGAACACCGCGGCATCAG 840
3359 CGGTCTTGACAAAAGAACCGCGCCCTGGCTGACAGCCGGAACACCGCGGCATCAG 3418
841 AGCAGCCGATGTCTGTTGCGCCAGTCAATAGCGGAATAGCTCTTCCACCCCAAGCGGCG 900
3419 AGCAGCCGATGTCTGTTGCGCCAGTCAATAGCGGAATAGCTCTTCCACCCCAAGCGGCG 3478
901 GAGAACCTGCGTGAATCCATCTTGTTCAATCATCGGAACGATCTCTCTCTCTT 960
3479 GAGAACCTGCGTGAATCCATCTTGTTCAATCATCGGAACGATCTCTCTCTCTT 3538
961 GATCAGATCTTGAATCCCTGCGCCATCAGATCCCTTGGCGGC 1001
3539 GATCAGATCTTGAATCCCTGCGCCATCAGATCCCTTGGCGGC 3579

RESULT 14

US-10-260-270-4
; Sequence 4, Application US/10260270
; Publication No. US20030143192A1
; GENERAL INFORMATION:
; APPLICANT: Salcedo, Theodora W.
; APPLICANT: Patel, Vikram, P.

APPLICANT: Nibbs, Robert J.B.
APPLICANT: Graham, Gerard J.
TITLE OF INVENTION: CHEMOKINE BETA-7
FILE REFERENCE: 1488.033000F
CURRENT APPLICATION NUMBER: US/10/260,270
CURRENT FILING DATE: 2002-10-01
PRIOR APPLICATION NUMBER: US/09/437,602
PRIOR FILING DATE: 1999-11-10
PRIOR APPLICATION NUMBER: US 60/107,801
PRIOR FILING DATE: 1998-11-10
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patent in version 3.0
SEQ ID NO 4
LENGTH: 3974
TYPE: DNA
ORGANISM: Homo sapiens
US-10-260-270-4

Query Match 100.0%; Score 1001; DB 13; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGAGGATCATCCAGCGCGCTCCCGAAGCGATTCGGAAGCCCACTTCATAG 60
Db 2579 CGCTGAGGATCATCCAGCGCGCTCCCGAAGCGATTCGGAAGCCCACTTCATAG 2638

QY 61 AAGCGCGGCTGGAATCGAATCTCGTATCGATGCGAGTTGGCGCTGCTTGTGCTGCTATT 120
Db 2639 AAGCGCGGCTGGAATCGAATCTCGTATCGATGCGAGTTGGCGCTGCTTGTGCTGCTATT 2698

QY 121 TGAACCCCGAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGCGGATGCGCT 180
Db 2699 TGAACCCCGAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGCGGATGCGCT 2758

QY 181 GCGATCGGAGCGCGATACGTAACGACAGAGAGCGGTGAGCCATTCGCGCCAA 240
Db 2759 GCGATCGGAGCGCGATACGTAACGACAGAGAGCGGTGAGCCATTCGCGCCAA 2818

QY 241 GCTCTTACGAATATCAGGCTAGCCAGCTATGCTGATAGCGTTCGCGCACACCA 300
Db 2819 GCTCTTACGAATATCAGGCTAGCCAGCTATGCTGATAGCGTTCGCGCACACCA 2878

QY 301 GCGGCGCAGTCGATGAATCAGAAAGCGGCAATTTCCACATGATATTCGGCAAG 360
Db 2879 GCGGCGCAGTCGATGAATCAGAAAGCGGCAATTTCCACATGATATTCGGCAAG 2938

QY 361 AGGCATCGCATGGTTCAGCAGAGATCTCGCGCTCGGCGATCGCGCTTCGAGCTGG 420
Db 2939 AGGCATCGCATGGTTCAGCAGAGATCTCGCGCTCGGCGATCGCGCTTCGAGCTGG 2998

QY 421 CGACAGTTCGGCTGGCGGAGCCCTGATGCTCTTCTCGATCATCTGATCGACAA 480
Db 2999 CGACAGTTCGGCTGGCGGAGCCCTGATGCTCTTCTCGATCATCTGATCGACAA 3058

QY 481 GACCGGCTTCATCGATGATGCTGCTCGATGCGATGCTTCTGCTGCTGCTGATG 540
Db 3059 GACCGGCTTCATCGATGATGCTGCTCGATGCGATGCTTCTGCTGCTGCTGATG 3118

QY 541 GCGAGTAGCGGATCAAGCGTATGAGCCCGCCATTCGATCAGCCATGATGATGATCT 600
Db 3119 GCGAGTAGCGGATCAAGCGTATGAGCCCGCCATTCGATCAGCCATGATGATGATCT 3178

QY 601 TCTCGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Db 3179 TCTCGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3238

QY 661 GCGAGTCCCTTCCCGCTTCAGTGAACGTCGACACAGTTCGCGAGGAGGAGGAGGAG 720
Db 3239 GCGAGTCCCTTCCCGCTTCAGTGAACGTCGACACAGTTCGCGAGGAGGAGGAGGAG 3298

QY 721 TGGCAGGACGATAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 3299 TGGCAGGACGATAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3358

QY 781 CGCTCTTGACAAAAGAACCGCGCGCTTGGCTGACAGCGGACGACGCGGCGATCAG 840
Db 3359 CGCTCTTGACAAAAGAACCGCGCGCTTGGCTGACAGCGGACGACGCGGCGATCAG 3418

QY 841 AGCAGCGGATTTGCTTGTGTGCGCCAGTCATAGCCGAATAGCCTTCTCCACCAAGCGCG 900
Db 3419 AGCAGCGGATTTGCTTGTGTGCGCCAGTCATAGCCGAATAGCCTTCTCCACCAAGCGCG 3478

QY 901 GAGAACTCGGTGCATCCATCTTGTTCATATGCGAAACGATCTCTCTCTCTCTCT 960
Db 3479 GAGAACTCGGTGCATCCATCTTGTTCATATGCGAAACGATCTCTCTCTCTCTCT 960

QY 961 GATCAGATCTTGATCCCTCGCGCCATCAGATCTTGGCGGC 1001
Db 3539 GATCAGATCTTGATCCCTCGCGCCATCAGATCTTGGCGGC 3579

RESULT 15
US-10-197-844-9
; Sequence 9, Application US/10197844
; Publication No. US20030166898A1
; GENERAL INFORMATION:
; APPLICANT: CHOPRA, ARVIND
; OLSEN, HENRIK
; GENIZ, STEVEN M.
; RUBIN, STEVEN M.
; TITLE OF INVENTION: MYELIN OLIGODENDROCYTE GLYCOPROTEIN-LIKE
; PROTEIN (MOGP) AND METHODS OF USE
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSER: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/197,844
; FILING DATE: 19-Jul-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/015,963
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 60/035,445
; FILING DATE: 30-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0520001/EKS/RCM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3974 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-10-197-844-9

Query Match 100.0%; Score 1001; DB 13; Length 3974;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGAGGATCATCCAGCGCGCTCCCGAAGCGATTCGGAAGCCCACTTCATAG 60
Db 2579 CGCTGAGGATCATCCAGCGCGCTCCCGAAGCGATTCGGAAGCCCACTTCATAG 2638

QY 61 AAGCGCGGCTGGAATCGAATCTCGTATCGATGCGAGTTGGCGCTGCTTGTGCTGCTATT 120
Db 2639 AAGCGCGGCTGGAATCGAATCTCGTATCGATGCGAGTTGGCGCTGCTTGTGCTGCTATT 2698

QY 121 TGAACCCCGAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGCGGATGCGCT 180
Db 2699 TGAACCCCGAGTCCCGCTCAGAGAACTCGTCAAGAGCGGATAGAGCGGATGCGCT 2758

QY 181 GCGATCGGAGCGCGATACGTAACGACAGAGAGCGGTGAGCCATTCGCGCCAA 240
Db 2759 GCGATCGGAGCGCGATACGTAACGACAGAGAGCGGTGAGCCATTCGCGCCAA 2818

QY 241 GCTCTTACGAATATCAGGCTAGCCAGCTATGCTGATAGCGTTCGCGCACACCA 300
Db 2819 GCTCTTACGAATATCAGGCTAGCCAGCTATGCTGATAGCGTTCGCGCACACCA 2878

QY 301 GCGGCGCAGTCGATGAATCAGAAAGCGGCAATTTCCACATGATATTCGGCAAG 360
Db 2879 GCGGCGCAGTCGATGAATCAGAAAGCGGCAATTTCCACATGATATTCGGCAAG 2938

QY 361 AGGCATCGCATGGTTCAGCAGAGATCTCGCGCTCGGCGATCGCGCTTCGAGCTGG 420
Db 2939 AGGCATCGCATGGTTCAGCAGAGATCTCGCGCTCGGCGATCGCGCTTCGAGCTGG 2998

QY 421 CGACAGTTCGGCTGGCGGAGCCCTGATGCTCTTCTCGATCATCTGATCGACAA 480
Db 2999 CGACAGTTCGGCTGGCGGAGCCCTGATGCTCTTCTCGATCATCTGATCGACAA 3058

QY 481 GACCGGCTTCATCGATGATGCTGCTCGATGCGATGCTTCTGCTGCTGCTGATG 540
Db 3059 GACCGGCTTCATCGATGATGCTGCTCGATGCGATGCTTCTGCTGCTGCTGATG 3118

QY 541 GCGAGTAGCGGATCAAGCGTATGAGCCCGCCATTCGATCAGCCATGATGATGATCT 600
Db 3119 GCGAGTAGCGGATCAAGCGTATGAGCCCGCCATTCGATCAGCCATGATGATGATCT 3178

QY 601 TCTCGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Db 3179 TCTCGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3238

QY 661 GCGAGTCCCTTCCCGCTTCAGTGAACGTCGACACAGTTCGCGAGGAGGAGGAGGAG 720
Db 3239 GCGAGTCCCTTCCCGCTTCAGTGAACGTCGACACAGTTCGCGAGGAGGAGGAGGAG 3298

QY 721 TGGCAGGACGATAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 3299 TGGCAGGACGATAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3358

```
Db 2579 CGCTGAGGATCATCCAGCGCGCTCCCGAAGAGATTCGAGAGCCCAACCTTTTCATAG 2638
QY 61 AAGCGCGCGTGGATCGAAATCTCGTATGCGAGTTGGCGTTCGCTTGGTTCGCTCAT 120
Db 2639 AAGCGCGCGTGGATCGAAATCTCGTATGCGAGTTGGCGTTCGCTTGGTTCGCTCAT 2698
QY 121 TCGAAGCCCGAGAGTCCCGCTCAGAAAGACTCGTCAAGAGGCGATAGAGGCGATGCGCT 180
Db 2699 TCGAAGCCCGAGAGTCCCGCTCAGAAAGACTCGTCAAGAGGCGATAGAGGCGATGCGCT 2758
QY 181 GCGATCGGAGCGGCGATACCGTAAAGACAGAGAGCGGTCAAGCCCATTCGCGGCCAA 240
Db 2759 GCGAATCGGAGCGGCGATACCGTAAAGACAGAGAGCGGTCAAGCCCATTCGCGGCCAA 2818
QY 241 GCTCTTCAGCAATATCAGGGTAGCAACGCTATGCTGATGAGGCTTCGCGGCCCA 300
Db 2819 GCTCTTCAGCAATATCAGGGTAGCAACGCTATGCTGATGAGGCTTCGCGGCCCA 2878
QY 301 GCGGCCACAGTTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATATTCGGCAAGC 360
Db 2879 GCGGCCACAGTTCGATGAATCCAGAAAGCGGCCATTTTCACCATGATATTCGGCAAGC 2938
QY 361 AGGCATCGCCATGGGTACAGAGAGATCCTCGCGTCGGGATGCGGCTTGAGCCTGG 420
Db 2939 AGGCATCGCCATGGGTACAGAGAGATCCTCGCGTCGGGATGCGGCTTGAGCCTGG 2998
QY 421 CGAACAGTTCGGCTGGCGGAGCGCCCTGATGCTCTTCGTCAGATCATCTGATCGACAA 480
Db 2999 CGAACAGTTCGGCTGGCGGAGCGCCCTGATGCTCTTCGTCAGATCATCTGATCGACAA 3058
QY 481 GACCGGCTTCATTCGAGTACGTGCTCGCTCGATGCGATGTTTCGTTGGTGGTTCGAATG 540
Db 3059 GACCGGCTTCATTCGAGTACGTGCTCGCTCGATGCGATGTTTCGTTGGTGGTTCGAATG 3118
QY 541 GCGAGTAGCGCGATCAAGCGTATGACGCGCGCGCATTTGATCAGCCATGATGATGATCTT 600
Db 3119 GCGAGTAGCGCGATCAAGCGTATGACGCGCGCGCATTTGATCAGCCATGATGATGATCTT 3178
QY 601 TCTGCGAGGAGCAAGGTGAGATCAAGAGATCCTGCGCGGACCTTCGCCCAATAGCA 660
Db 3179 TCTGCGAGGAGCAAGGTGAGATCAAGAGATCCTGCGCGGACCTTCGCCCAATAGCA 3238
QY 661 GCCAGTCCCTTCGCGCTTCAGTGAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 720
Db 3239 GCCAGTCCCTTCGCGCTTCAGTGAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 3298
QY 721 TGGCCAGCCAGATAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 3299 TGGCCAGCCAGATAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3358
QY 781 CGGTCTTGACAAAGAACCGGGCGCGCTGCGCTGAGCAGCGGAAACCGCGGATCAG 840
Db 3359 CGGTCTTGACAAAGAACCGGGCGCGCTGCGCTGAGCAGCGGAAACCGCGGATCAG 3418
QY 841 AGCAGCGATGTGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 3419 AGCAGCGATGTGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3478
QY 901 GAGAACCTGCTGCAATCCATCTTGTTCATATGCGAAACGATCCCTCATCTCTCTCT 960
Db 3479 GAGAACCTGCTGCAATCCATCTTGTTCATATGCGAAACGATCCCTCATCTCTCTCT 3538
QY 961 GATCAGATCTTGATCCCTGCGCATCAGATCCCTGCGGC 1001
Db 3539 GATCAGATCTTGATCCCTGCGCATCAGATCCCTGCGGC 3579
```

Search completed: February 15, 2004, 22:26:53
Job time : 425.484 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 21:50:21 ; Search time 2636.53 Seconds
(without alignments)
9227.587 Million cell updates/sec

Title: US-09-921-143-36_COPY_4000_5000

Perfect score: 1001

Sequence: 1 ccttgagatcaccagcc.....gcatcgatccttggcgcg 1001

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size : 0

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_estin:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htrc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htrc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gas_hum:*

18: em_gas_inv:*

19: em_gas_pln:*

20: em_gas_vrt:*

21: em_gas_fun:*

22: em_gas_mam:*

23: em_gas_mus:*

24: em_gas_pro:*

25: em_gas_rtd:*

26: em_gas_phg:*

27: em_gas_vrl:*

28: gb_gaal:*

29: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	568	56.7	784	28	AQ876152 V133H12 m
2	568	56.7	850	28	AQ875052 V120F5 mt
3	520	51.9	798	28	AQ876139 V133G10 m
4	508	50.7	509	28	AQ447775 mgxb0011E

5	508	50.7	790	28	AQ876067	AQ876067 V132H11 m
6	507	50.6	789	28	AQ361914	AQ361914 mgxb00005K
7	499	49.9	499	28	AQ398880	AQ398880 mgxb00006P
8	499	49.9	592	28	AQ449182	AQ449182 mgxb00023E
9	493	49.3	637	28	AQ447230	AQ447230 mgxb00006G
10	493	49.3	861	28	AQ875040	AQ875040 V12084 mt
11	489	48.9	799	28	AQ876220	AQ876220 V15267 mt
12	481	48.1	801	28	AQ875907	AQ875907 V130H3 mt
13	479	47.9	783	28	AQ876024	AQ876024 V132C9 mt
14	468	46.8	561	28	AQ447874	AQ447874 mgxb00012I
15	468	46.8	620	28	AQ398387	AQ398387 mgxb00009J
16	468	46.8	789	28	AQ876017	AQ876017 V132C12 m
17	459	45.9	779	28	AQ875906	AQ875906 V13082 mt
18	459	45.9	791	28	AQ876019	AQ876019 V132C3 mt
19	459	45.9	857	28	AQ875013	AQ875013 V12088 mt
20	457	45.7	791	28	AQ876015	AQ876015 V132C10 m
21	454	45.4	498	28	AQ398730	AQ398730 mgxb00009G
22	445	44.5	635	28	AQ447643	AQ447643 mgxb00008M
23	445	44.5	656	28	AQ447140	AQ447140 mgxb00005I
24	438	43.8	797	28	AQ876297	AQ876297 V153G3 mt
25	435	43.5	791	28	AQ876201	AQ876201 V15282 mt
26	429	42.9	473	28	AQ397768	AQ397768 mgxb00001J
27	428	42.8	798	28	AQ875840	AQ875840 V130B11 m
28	410	41.0	795	28	AQ876130	AQ876130 V133F12 m
29	398	39.8	790	28	AQ875904	AQ875904 V130H11 m
30	397	39.7	860	28	AQ875029	AQ875029 V120D2 mt
31	390	39.0	762	28	AQ873181	AQ873181 V61A6 mTn
32	387	38.7	815	28	AQ876134	AQ876134 V133P5 mt
33	382	38.2	748	28	AQ501530	AQ501530 V20F6 mTn
34	381	38.1	382	28	AQ448898	AQ448898 mgxb00022E
35	377	37.7	540	28	AQ874387	AQ874387 V107H5 mt
36	371	37.1	422	28	BH912281	BH912281 PST277.se
37	366	36.6	642	28	AQ874213	AQ874213 V104H2 mt
38	366	36.6	801	28	AQ876288	AQ876288 V153F3 mt
39	363	36.3	366	28	AQ399385	AQ399385 mgxb0010J
40	357	35.7	704	28	AQ874178	AQ874178 V104C5 mt
41	355	35.5	804	28	AQ361943	AQ361943 mgxb00005M
42	353	35.3	500	12	BM056700	BM056700 2107-31 h
43	351	35.1	734	28	AQ873679	AQ873679 V73F2 mTn
44	349	34.9	779	28	AQ876036	AQ876036 V132E1 mt
45	348	34.8	705	28	AQ500916	AQ500916 V30D10 mt

ALIGNMENTS

RESULT 1
AQ876152
LOCUS
DEFINITION
V133H12 mTn-3xHA/lacZ Insertion Library, strain Y2278 Saccharomyces cerevisiae genomic 5', genomic survey sequence.
ACCESSION
AQ876152.1 GI:6288396
VERSION
AQ876152.1
KEYWORDS
GSS.
SOURCE
Saccharomyces cerevisiae (baker's yeast)
ORGANISM
Saccharomyces cerevisiae
REFERENCE
1 (bases 1 to 784)
AUTHORS
Roos-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A., deStagis, S.A., Cheung, K.-H., Sheehan, A., Symoniat, D., Jansen, R., Umansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R., Hager, K., Miller, P., Roeder, G.S. and Snyder, M.
TITLE
Large-Scale Analysis of the Yeast Genome by Transposon Tagging and Gene Disruption
JOURNAL
Unpublished
COMMENT
Contact: Kumar A
Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
Yale University
P.O. Box 208103, New Haven, CT 06520-8103, USA
Tel: 203 432 9949
Fax: 203 432 6161
Email: anuj.kumar@yale.edu

te of mTn-3xHA/lacZ insertion.
Seq primer: GGCTTCTTTCTTGGAGTAC
Class: transposon-tagged.

FEATURES

Source
Location/Qualifiers

1..784
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/strain="Y2278 - S288C background, cir(0) rho(0)"
/db_xref="taxon:4932"
/lab_host="E. coli"
/clone_lib="mTn-3xHA/lacZ Insertion Library, strain Y2278"
/note="Vector: pHS6-Sal; A yeast genomic DNA library without 2 micron or mitochondrial DNA was prepared in pHS6-Sal; genomic DNA was size-fractionated (DNA of roughly 2-3 kb in length) prior to cloning. This library was subsequently mutagenized with a mTn-3xHA/lacZ minitransposon containing lacZ, URA3, and tet resistance."

BASE COUNT 166 a 226 c 218 g 172 t 2 others
ORIGIN

Query Match 56.7%; Score 568; DB 28; Length 784;
Best Local Similarity 100.0%; Pred. No. 5.6e-287;
Matches 568; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGAGGATCATCCAGCGCGTCCGGAAACGATTCGAGCCCAACCTTTCATAG 60
DB 195 CGCTGAGGATCATCCAGCGCGTCCGGAAACGATTCGAGCCCAACCTTTCATAG 254
QY 61 AAGGCGCGGTGGAATCTCGTATGCGAGTGGCGTGGCTTGGTGGTCAATT 120
DB 255 AAGGCGCGGTGGAATCTCGTATGCGAGTGGCGTGGCTTGGTGGTCAATT 314
QY 121 TCGAACCCAGAGTCCCGTCAAGAACTGTCAGAGGCGATAGAGCGATGCGCT 180
DB 315 TCGAACCCAGAGTCCCGTCAAGAACTGTCAGAGGCGATAGAGCGATGCGCT 374
QY 181 GCGAATCGGAGCGCGGATACCGTAAAGCAGAGGAGCGGTCCAGCCCATTCGCGGCAA 240
DB 375 GCGAATCGGAGCGCGGATACCGTAAAGCAGAGGAGCGGTCCAGCCCATTCGCGGCAA 434
QY 241 GCTCTTCAGCATATACGGGTAGCAACGCTATGCTGATAGCGGTTCGCGCAACCA 300
DB 435 GCTCTTCAGCATATACGGGTAGCAACGCTATGCTGATAGCGGTTCGCGCAACCA 494
QY 301 GCGGCGCAAGTCGATGAATCCAGAAAGCGGCCATTTTCCACCATGATATTCGCGAAGC 360
DB 495 GCGGCGCAAGTCGATGAATCCAGAAAGCGGCCATTTTCCACCATGATATTCGCGAAGC 554
QY 361 AGGCATCGCCATGGGTACGACAGATCCTCGCGTGGGATCGCGCTTGGAGCCTGG 420
DB 555 AGGCATCGCCATGGGTACGACAGATCCTCGCGTGGGATCGCGCTTGGAGCCTGG 614
QY 421 CGAACAGTTCGGTGGCGGAGCCCTGATGCTCTTCGTCAGATCATCTGATCGCAAA 480
DB 615 CGAACAGTTCGGTGGCGGAGCCCTGATGCTCTTCGTCAGATCATCTGATCGCAAA 674
QY 481 GACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTCAATG 540
DB 675 GACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTCAATG 734
QY 541 GCGAGGTAGCGGATCAAGCGTATCGAG 568
DB 735 GCGAGGTAGCGGATCAAGCGTATCGAG 762

RESULT 2
A0875052
LOCUS 850 bp DNA linear GSS 08-NOV-1999
DEFINITION V120F5 mTn-3xHA/lacZ Insertion Library, strain Y2278 Saccharomyces cerevisiae genomic 5', genomic survey sequence.
ACCESSION A0875052
VERSION A0875052.1 GI:6287296

KEYWORDS

SOURCE ORGANISM
Saccharomyces cerevisiae (baker's yeast)

REFERENCE

AUTHORS

1 (bases 1 to 850)
Ross-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A., desEtages, S.A., Cheung, K.-H., Sheehan, A., Symoniatitis, D., Jansen, R., Umansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R., Hager, K., Miller, P., Roeder, G.S. and Snyder, M.

TITLE

Large-Scale Analysis of the Yeast Genome by Transposon Tagging and Gene Disruption

JOURNAL

COMMENT

Contact: Kumar A
Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
Yale University
P.O. Box 208103, New Haven, CT 06520-8103, USA
Tel: 203 432 9949
Fax: 203 432 6161
Email: anuj.kumar@yale.edu
te of mTn-3xHA/lacZ insertion.
Seq primer: GGCTTCTTTCTTGGAGTAC
Class: transposon-tagged.

FEATURES

Location/Qualifiers

1..850
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/strain="Y2278 - S288C background, cir(0) rho(0)"
/db_xref="taxon:4932"
/lab_host="E. coli"
/clone_lib="mTn-3xHA/lacZ Insertion Library, strain Y2278"
/note="Vector: pHS6-Sal; A yeast genomic DNA library without 2 micron or mitochondrial DNA was prepared in pHS6-Sal; genomic DNA was size-fractionated (DNA of roughly 2-3 kb in length) prior to cloning. This library was subsequently mutagenized with a mTn-3xHA/lacZ minitransposon containing lacZ, URA3, and tet resistance."

BASE COUNT 183 a 247 c 246 g 172 t 2 others
ORIGIN

Query Match 56.7%; Score 568; DB 28; Length 850;
Best Local Similarity 100.0%; Pred. No. 5.7e-287;
Matches 568; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGAGGATCATCCAGCGCGTCCGGAAACGATTCGAGCCCAACCTTTCATAG 60
DB 179 CGCTGAGGATCATCCAGCGCGTCCGGAAACGATTCGAGCCCAACCTTTCATAG 238
QY 61 AAGGCGCGGTGGAATCTCGTATGCGAGTGGCGTGGCTTGGTGGTCAATT 120
DB 239 AAGGCGCGGTGGAATCTCGTATGCGAGTGGCGTGGCTTGGTGGTCAATT 298
QY 121 TCGAACCCAGAGTCCCGTCAAGAACTGTCAGAGGCGATAGAGCGATGCGCT 180
DB 299 TCGAACCCAGAGTCCCGTCAAGAACTGTCAGAGGCGATAGAGCGATGCGCT 358
QY 181 GCGAATCGGAGCGCGGATACCGTAAAGCAGAGGAGCGGTCCAGCCCATTCGCGGCAA 240
DB 359 GCGAATCGGAGCGCGGATACCGTAAAGCAGAGGAGCGGTCCAGCCCATTCGCGGCAA 418
QY 241 GCTCTTCAGCATATACGGGTAGCAACGCTATGCTGATAGCGGTTCGCGCAACCA 300
DB 419 GCTCTTCAGCATATACGGGTAGCAACGCTATGCTGATAGCGGTTCGCGCAACCA 478
QY 301 GCGGCGCAAGTCGATGAATCCAGAAAGCGGCCATTTTCCACCATGATATTCGCGAAGC 360
DB 479 GCGGCGCAAGTCGATGAATCCAGAAAGCGGCCATTTTCCACCATGATATTCGCGAAGC 538
QY 361 AGGCATCGCCATGGGTACGACAGATCCTCGCGTGGGATCGCGCTTGGAGCCTGG 420
DB 539 AGGCATCGCCATGGGTACGACAGATCCTCGCGTGGGATCGCGCTTGGAGCCTGG 598

QY 421 CGAAGAGTTCCGGTGGCGAGCCCTGATGCTCTTCCTCCAGATCATCTGATCGCAA 480
 Db 599 CGAAGAGTTCCGGTGGCGAGCCCTGATGCTCTTCCTCCAGATCATCTGATCGCAA 658
 QY 481 GACCGGCTTCCATCCGAGTACGTGCTCGATCGGATGTTTCCTGTTGGTGGTGAATG 540
 Db 659 GACCGGCTTCCATCCGAGTACGTGCTCGATCGGATGTTTCCTGTTGGTGGTGAATG 718
 QY 541 GGCAGTAGCCGGATCAAGCGTATGCG 568
 Db 719 GGCAGTAGCCGGATCAAGCGTATGCG 746

RESULT 3
 LOCUS AQB76139 798 bp DNA linear GSS 08-NOV-1999
 DEFINITION V133G10 mTn-3xHA/lacZ Insertion Library, strain Y2278 Saccharomyces cerevisiae genomic 5', genomic survey sequence.

ACCESSION AQB76139
 VERSION AQB76139.1 GI:6288383
 KEYWORDS GSS.
 SOURCE Saccharomyces cerevisiae (baker's yeast)
 ORGANISM Saccharomyces cerevisiae
 Eukaryota; Fungi; Ascomycota; Saccharomycetaceae; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

REFERENCE 1 (bases 1 to 798)
 AUTHORS Ross-Macdonald, P., Roemer, T., Coelho, P. S. R., Agarwal, S., Kumar, A., deStages, S. A., Cheung, K. H., Sheehan, A., Symoniat, D., Jansen, R., Umansky, L., Heidman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R., Hager, K., Miller, P., Roder, G. S. and Snyder, M.
 TITLE Large-Scale Analysis of the Yeast Genome by Transposon Tagging and Gene Disruption
 JOURNAL Unpublished
 COMMENT Contact: Kumar A
 Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
 Yale University
 P.O. Box 208103, New Haven, CT 06520-8103, USA
 Tel: 203 432 9949
 Fax: 203 432 6161
 Email: anuj.kumar@yale.edu
 te of mTn-3xHA/lacZ insertion.
 Seq primer: GGCCTCTTCTTGTGGAAGTAC
 Class: transposon-tagged

FEATURES
 source
 Location/Qualifiers
 1..798
 /organism="Saccharomyces cerevisiae"
 /mol_type="genomic DNA"
 /strain="Y2278 - S288C background, cir(0) rho(0)"
 /db_xref="taxon:4932"
 /lab_host="E. coli"
 /clone_lib="mTn-3xHA/lacZ Insertion Library, strain Y2278"
 /note="Vector: pHS56-Sal; A yeast genomic DNA library without 2 micron or mitochondrial DNA was prepared in pHS56-Sal; genomic DNA was size-fractionated (DNA of roughly 2-3 kb in length) prior to cloning. This library was subsequently mutagenized with a mTn-3xHA/lacZ minitransposon containing lacZ, URA3, and tet resistance."

BASE COUNT 176 a 230 c 226 g 164 t 2 others
 ORIGIN
 Query Match 51.9%; Score 520; DB 28; Length 798;
 Best Local Similarity 99.8%; Pred No. 9.3e-262;
 Matches 570; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCTGGAGGATCATCCAGCGCGTCCCGAAACAGATTCCGAAGCCCACTTTCATAG 60
 Db 194 CGCTGGAGGATCATCCAGCGCGTCCCGAAACAGATTCCGAAGCCCACTTTCATAG 253
 QY 61 AAGGCGCGGTGGATCGAATCTCGTATGCGAGTTGGCGTGGTGGTGGTGGTGGT 120
 Db 254 AAGGCGCGGTGGATCGAATCTCGTATGCGAGTTGGCGTGGTGGTGGTGGTGGT 313

QY 121 TCGAAGCCCAAGTCCCGCTCAGAGAACTCGTCAAGAGGCGATAGAGGCGATCGCT 180
 Db 314 TCGAAGCCCAAGTCCCGCTCAGAGAACTCGTCAAGAGGCGATAGAGGCGATCGCT 373
 QY 181 GCGAATCGGAGCGGCGATACCGTAAAGCAGAGGAGCGGTGAGCCCATTCGCCGCCAA 240
 Db 374 GCGAATCGGAGCGGCGATACCGTAAAGCAGAGGAGCGGTGAGCCCATTCGCCGCCAA 433
 QY 241 GCTCTTACGAATATCACCGGTAGCCACGCTATGCTCTGATAGCGGTTCGCCACACCCA 300
 Db 434 GCTCTTACGAATATCACCGGTAGCCACGCTATGCTCTGATAGCGGTTCGCCACACCCA 493
 QY 301 GCCGGCCACAGTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATTCGGCAAGC 360
 Db 494 GCCGGCCACAGTCGATGAATCCAGAAAGCGGCCATTTCCACCATGATTCGGCAAGC 553
 QY 361 AGGCATCGCCATCGGTCACGACGAGATCTTCGCCGTCGGGCGATCGCGGCTTGAGCCTGG 420
 Db 554 AGGCATCGCCATCGGTCACGACGAGATCTTCGCCGTCGGGCGATCGCGGCTTGAGCCTGG 613
 QY 421 CGAAGAGTTCCGGTGGCGAGCCCTGATGCTCTTCGTCAGATCATCTGATCGCAA 480
 Db 614 CGAAGAGTTCCGGTGGCGAGCCCTGATGCTCTTCGTCAGATCATCTGATCGCAA 673
 QY 481 GACCGGCTTCCATCCGAGTACGTGCTCGATCGGATGTTTCGCTTGGTGGTGAATG 540
 Db 674 GACCGGCTTCCATCCGAGTACGTGCTCGATCGGATGTTTCGCTTGGTGGTGAATG 733
 QY 541 GGCAGTAGCCGGATCAAGCGTATGCGCG 571
 Db 734 GGCAGTAGCCGGATCAAGCGTATGCGCG 764

RESULT 4
 LOCUS AQ447775/c 509 bp DNA linear GSS 08-APR-1999
 DEFINITION mgxb0011E13f CUGI Rice Blast BAC Library Magnaporthe grisea genomic clone mgxb0011E13f, genomic survey sequence.

ACCESSION AQ447775
 VERSION AQ447775.1 GI:4576912
 KEYWORDS GSS.
 SOURCE Magnaporthe grisea (anamorph: Pyricularia grisea)
 ORGANISM Magnaporthe grisea
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.

REFERENCE 1 (bases 1 to 509)
 AUTHORS Yu, Y., Zhu, H., Boyd, C. A., Gaudette, B., Gayle, A., Kingsbury, R., Phillips, K., Sasinowski, M., Wing, R. A. and Dean, R. A.
 TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome
 JOURNAL Unpublished
 COMMENT Contact: Dean RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson University, Clemson, SC 29634
 Tel: 864 656 5737
 Fax: 864 656 4293
 Email: rdean@clemson.edu
 Seq primer: TAATACGACTCATATAGG
 Class: BAC ends
 High quality sequence stop: 342.

FEATURES
 source
 Location/Qualifiers
 1..509
 /organism="Magnaporthe grisea"
 /mol_type="genomic DNA"
 /strain="70-15"
 /db_xref="taxon:148305"
 /clone="mgxb0011E13f"
 /tissue_type="Protoplasts"
 /lab_host="E. coli DH10B"
 /clone_lib="CUGI Rice Blast BAC Library"
 /note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
 Rice blast is one of the most devastating fungal diseases


```

clone mgxb0005K01f, genomic survey sequence.
ACCESSION AQ361914 GI:42111753
VERSION AQ361914.1
KEYWORDS GSS.
SOURCE Magnaporthe grisea (anamorph: Pyricularia grisea)
ORGANISM Magnaporthe grisea
REFERENCE Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
AUTHORS Sordariomycetes incertae sedis; Magnaportheae; Magnaporthe.
1 (bases 1 to 789)
Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
JOURNAL Genome
COMMENT Unpublished
CONTACT: Dean RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737
Fax: 864 656 4293
Email: rdean@clemson.edu
Seq primer: TAATACGACTCACTATAGG
Class: BAC ends
High quality sequence start: 41
High quality sequence stop: 392.
Location/Qualifiers
1..789
/organism="Magnaporthe grisea"
/mol_type="genomic DNA"
/strain="70-15"
/db_xref="taxon:148305"
/clone="mgxb0005K01f"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/clone_lib="CUGI Rice Blast BAC Library"
/note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
Rice blast is one of the most devastating fungal diseases
of rice world wide. It is a filamentous ascomycete with
a haploid genome (n=7) of approximately 40 Mbp. Rice
blast is an important model fungal pathogen for studying
numerous aspects of the fungal-host interaction. In
order to facilitate genome wide analysis, a BAC library
containing 9216 clones with an average insert size of 130
kbp was constructed. This library represents greater
than 25X genome coverage. High density colony filters
are available upon request."
BASE COUNT 160 a 211 c 234 g 181 t 3 others
ORIGIN
Query Match 50.6%; Score 507; DB 28; Length 789;
Best Local Similarity 99.8%; Pred. No. 6.3e-255;
Matches 557; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 180 TCGGAATCGGAGCGCGATACCGTAAAGCAGCAGGAGCGGTGAGCCCATTCGCGGCCA 239
DB 599 TCGGAATCGGAGCGCGCGATACCGTAAAGCAGCAGGAGCGGTGAGCCCATTCGCGGCCA 540
QY 240 AGCTCTTCAGCAATATCAGCGGTAGCCAAACGCTATGCTCTGATAGCGGTGCGCACACC 299
DB 539 AGCTCTTCAGCAATATCAGCGGTAGCCAAACGCTATGCTCTGATAGCGGTGCGCACACC 480
QY 300 AGCCGGCCACACTGATGATTCACAAAAGCGGCCATTTCCACCATGATATTCGGCAAG 359
DB 479 AGCCGGCCACACTGATGATTCACAAAAGCGGCCATTTCCACCATGATATTCGGCAAG 420
QY 360 CAGGCATCCCATGGGTACGACGAGATCCTCGCGGTGCGGCATGCGCGCTTGAGCTTG 419
DB 419 CAGGCATCCCATGGGTACGACGAGATCCTCGCGGTGCGGCATGCGCGCTTGAGCTTG 360
QY 420 GCGAACAGTTCGCTGGCGGAGCCCTTGATGCTCTTCGTCGAGATCATCTCATGACACA 479
DB 359 GCGAACAGTTCGCTGGCGGAGCCCTTGATGCTCTTCGTCGAGATCATCTCATGACACA 300
clone mgxb0005K01f, genomic survey sequence.
QY 480 AGACCGGCTTCATCCGATGATGCTCGCTCGATCGATGATGCTCGCTTGTGTGTCGAAT 539
DB 299 AGACCGGCTTCATCCGATGATGCTCGCTCGATCGATGATGCTCGCTTGTGTGTCGAAT 240
QY 540 GGGCAGTAGCCGATCAACGCTATGACGCGCGCATTCGATCGATCGATCGATCGAT 599
DB 239 GGGCAGTAGCCGATCAACGCTATGACGCGCGCATTCGATCGATCGATCGATCGAT 180
QY 600 TTCTCGCAGGAGCAGAGTGAGATGACAGAGATCTGCGCGGCACTTCGCCCATAGC 659
DB 179 TTCTCGCAGGAGCAGAGTGAGATGACAGAGATCTGCGCGGCACTTCGCCCATAGC 120
QY 660 AGCCAGTCCCTTCGCTTCAGTGACAAAGTCGAGCAGCAGCTGCGCAAGAACGCCGTC 719
DB 119 AGCCAGTCCCTTCGCTTCAGTGACAAAGTCGAGCAGCAGCTGCGCAAGAACGCCGTC 60
QY 720 GTGGCCAGCCAGCATAGC 737
DB 59 GTGGCCAGCCAGCATAGC 42
RESULT 7
AQ398880/c
LOCUS 499 bp DNA linear GSS 06-MAR-1999
DEFINITION mgxb0006P12f CUGI Rice Blast BAC Library Magnaporthe grisea genomic
clone mgxb0006P12f, genomic survey sequence.
ACCESSION AQ398880
VERSION AQ398880.1 GI:4369907
KEYWORDS GSS.
SOURCE Magnaporthe grisea (anamorph: Pyricularia grisea)
ORGANISM Magnaporthe grisea
REFERENCE Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
AUTHORS Sordariomycetes incertae sedis; Magnaportheae; Magnaporthe.
1 (bases 1 to 499)
Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
JOURNAL Genome
COMMENT Unpublished
CONTACT: Dean RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737
Fax: 864 656 4293
Email: rdean@clemson.edu
Seq primer: TAATACGACTCACTATAGG
Class: BAC ends
High quality sequence stop: 422.
Location/Qualifiers
1..499
/organism="Magnaporthe grisea"
/mol_type="genomic DNA"
/strain="70-15"
/db_xref="taxon:148305"
/clone="mgxb0006P12f"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/clone_lib="CUGI Rice Blast BAC Library"
/note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
Rice blast is one of the most devastating fungal diseases
of rice world wide. It is a filamentous ascomycete with
a haploid genome (n=7) of approximately 40 Mbp. Rice
blast is an important model fungal pathogen for studying
numerous aspects of the fungal-host interaction. In
order to facilitate genome wide analysis, a BAC library
containing 9216 clones with an average insert size of 130
kbp was constructed. This library represents greater
than 25X genome coverage. High density colony filters
are available upon request."
BASE COUNT 94 a 142 c 157 g 106 t
ORIGIN

```

Query Match 49.9%; Score 499; DB 28; Length 499;
 Best Local Similarity 100.0%; Pred. No. 9.3e-251;
 Matches 499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 GCCTGCGAATCGGAGCGCGATACCGTAAGACAGAGAGCGGTACGCCCATTCGCC 235
 Db 499 GCCTGCGAATCGGAGCGCGATACCGTAAGACAGAGAGCGGTACGCCCATTCGCC 440

QY 236 GCGAAGCTCTTACGCAATATCAAGGTAGCCAAAGCGTATGCTGATAGCGGTCCGCCAC 295
 Db 439 GCGAAGCTCTTACGCAATATCAAGGTAGCCAAAGCGTATGCTGATAGCGGTCCGCCAC 380

QY 296 ACCGAGCGCGACAGTCGATGATCCAGAAAGCGGCCATTTCCACATGATATTCGG 355
 Db 379 ACCGAGCGCGACAGTCGATGATCCAGAAAGCGGCCATTTCCACATGATATTCGG 320

QY 356 CAAGAGGCGATCGCCATGGGTACGACGAGAGATCCTCGCGTCGGCGATGCGCGCTTCGAG 415
 Db 319 CAAGAGGCGATCGCCATGGGTACGACGAGAGATCCTCGCGTCGGCGATGCGCGCTTCGAG 260

QY 416 CTGCGGACAGCTTCGGCTGGCGAGCGCCCTGATGCTCTTCGTCAGATCATCTGATC 475
 Db 259 CTGCGGACAGCTTCGGCTGGCGAGCGCCCTGATGCTCTTCGTCAGATCATCTGATC 200

QY 476 GACAAGACCGGCTTCATCCGAGTACGTCCTCGCTCGATGCGATGCTTCGCTTGGTGTC 535
 Db 199 GACAAGACCGGCTTCATCCGAGTACGTCCTCGCTCGATGCGATGCTTCGCTTGGTGTC 140

QY 536 GAATGGGAGGTAGCGCGATCAAGGTATGAGCGCGCGATGCGCGCGATGCGCGATGCGA 595
 Db 139 GAATGGGAGGTAGCGCGATCAAGGTATGAGCGCGCGATGCGCGCGCGATGCGCGATGCGA 80

QY 596 TACTTCTTCGCGAGGAGCAAGGTGAGTACAGAGAGATCCTCGCGCGCGATGCGCGCGA 655
 Db 79 TACTTCTTCGCGAGGAGCAAGGTGAGTACAGAGAGATCCTCGCGCGCGATGCGCGCGA 20

QY 656 TAGCGCCAGTCCCTTCCC 674
 Db 19 TAGCGCCAGTCCCTTCCC 1

RESULT 8
 AQ449162/c
 LOCUS
 DEFINITION mgxb00023E21f CUGI Rice Blast BAC Library Magnaporthe grisea genomic clone mgxb00023E21f, genomic survey sequence.
 ACCESSION AQ449162
 VERSION
 KEYWORDS GSS
 SOURCE Magnaporthe grisea (anamorph: Pyricularia grisea)
 ORGANISM
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Magnaporthe grisea
 1 (bases 1 to 592)
 Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R., Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
 A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome
 JOURNAL Unpublished
 COMMENT Contact: Dean RA
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson University, Clemson, SC 29634
 Tel: 864 656 5737
 Fax: 864 656 4293
 Email: rdean@clemson.edu
 Seq primer: TAATACGACTCACTATAGGG
 Class: BAC ends
 High quality sequence stop: 395.
 Location/Qualifiers
 1..592
 /organism="Magnaporthe grisea"
 /mol_type="genomic DNA"

FEATURES
 source

/strain="70-15"
 /db_xref="taxon:148305"
 /clone="mgxb00023E21f"
 /issue_type="Protoplasts"
 /lab_host="E. coli DH10B"
 /clone_lib="CUGI Rice Blast BAC Library"
 /notes="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
 Rice blast is one of the most devastating fungal diseases of rice world wide. It is a filamentous ascomycete with a haploid genome (n=7) of approximately 40 Mbp. Rice blast is an important model fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25X genome coverage. High density colony filters are available upon request."

BASE COUNT 118 a 166 c 172 g 136 t
 ORIGIN

Query Match 49.9%; Score 499; DB 28; Length 592;
 Best Local Similarity 100.0%; Pred. No. 9.6e-251;
 Matches 499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 CGTCGCTTGGTGGTCAATTCGAAACCCAGAGTCCCGCTCAGAAGAACTCGTCAAGAGG 161
 Db 544 CGTCGCTTGGTGGTCAATTCGAAACCCAGAGTCCCGCTCAGAAGAACTCGTCAAGAGG 485

QY 162 CGATAGAAGCGGATCGGCTGCGAATCGGAGCGCGATACCGTAAAGACAGAGGAGCGG 221
 Db 484 CGATAGAAGCGGATCGGCTGCGAATCGGAGCGCGATACCGTAAAGACAGAGGAGCGG 425

QY 222 TGAGCCCATTCGCGCCCAAGCTCTTCAGCAATATCACGGGTAGCCAAACGCTATGTCCTGA 281
 Db 424 TCAGCCCATTCGCGCCCAAGCTCTTCAGCAATATCACGGGTAGCCAAACGCTATGTCCTGA 365

QY 282 TAGCGGTCCGCCACACCCAGCGGCACAGTCGATGATATCCAGAAAGCGGCATTTCC 341
 Db 364 TAGCGGTCCGCCACACCCAGCGGCACAGTCGATGATATCCAGAAAGCGGCATTTCC 305

QY 342 ACCATGATATTCGCGAAGCAGGATCGCCATCGGTCACGACGAGATCCTTCGCGCGCGG 401
 Db 304 ACCATGATATTCGCGAAGCAGGATCGCCATCGGTCACGACGAGATCCTTCGCGCGCGG 245

QY 402 ATGCGCGCTTCGAGCTGCGCAACAGTTCGGTTCGCGCGAGCCCTGATGCTCTTCGTC 461
 Db 244 ATGCGCGCTTCGAGCTGCGCAACAGTTCGGTTCGCGCGAGCCCTGATGCTCTTCGTC 185

QY 462 AGATCATCTCTGATCGACAAAGACCGGCTTCCATCCGAGTACGTCGCTCGATGCGATGT 521
 Db 184 AGATCATCTCTGATCGACAAAGACCGGCTTCCATCCGAGTACGTCGCTCGATGCGATGT 125

QY 522 TTGCGCTTGGTTCGAATGGGAGGTAGCGGATCAAGCGTATGAGCGCGCGCATTCGA 581
 Db 124 TTGCGCTTGGTTCGAATGGGAGGTAGCGGATCAAGCGTATGAGCGCGCGCATTCGA 65

QY 582 TCAGCCATGATGATCTT 600
 Db 64 TCAGCCATGATGATCTT 46

RESULT 9
 AQ447230/c
 LOCUS
 DEFINITION mgxb00060O09f CUGI Rice Blast BAC Library Magnaporthe grisea genomic clone mgxb00060O09f, genomic survey sequence.
 ACCESSION AQ447230
 VERSION
 KEYWORDS GSS
 SOURCE Magnaporthe grisea (anamorph: Pyricularia grisea)
 ORGANISM
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Magnaporthe grisea
 637 bp DNA linear GSS 08-APR-1999
 clone mgxb00060O09f, genomic survey sequence.
 AQ447230.1 GI:4576367
 GSS
 Magnaporthe grisea (anamorph: Pyricularia grisea)
 Magnaporthe grisea
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetes incertae sedis; Magnaportheae; Magnaporthe.

REFERENCE 1 (bases 1 to 637)
 AUTHORS Yu, Y., Zhu, H., Boyd, C. A., Gaudette, B., Gayle, A., Kingsbury, R., Phillips, K., Sasinski, M., Wing, R. A., and Dean, R. A.
 TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome
 JOURNAL Unpublished
 COMMENT Contact: Dean RA
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson University, Clemson, SC 29634
 Tel: 864 656 5737
 Fax: 864 656 4293
 Email: rdean@clemson.edu
 Seq primer: TAATACGACTCACTATAGGG
 Class: BAC ends
 High quality sequence stop: 291.
 FEATURES Location/Qualifiers
 source 1..637
 /organism="Magnaporthe grisea"
 /mol_type="genomic DNA"
 /strain="70-15"
 /db_xref="taxon:148305"
 /clone="mgxb0006009f"
 /tissue_type="Protoplasts"
 /lab_host="E. coli DH10B"
 /clone_lib="CUGI Rice Blast BAC Library"
 /note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII; Rice blast is one of the most devastating fungal diseases of rice world wide. It is a filamentous ascomycete with a haploid genome (n=7) of approximately 40 Mbp. Rice blast is an important model fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25X genome coverage. High density colony filters are available upon request."
 BASE COUNT 127 a 170 c 187 g 152 t 1 others

Query Match 49.3%; Score 493; DB 28; Length 637;
 Best Local Similarity 99.8%; Pred. No. 1.4e-247;
 Matches 543; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 102 CGTCGCTTGTGGTCTGTTGCAACCCAGAGTCCCGCTCAGAAGAACTCGTCAAGAGG 161
 DB 544 CGTCGCTTGTGGTCTGTTGCAACCCAGAGTCCCGCTCAGAAGAACTCGTCAAGAGG 485
 QY 162 CGATAGAAGCGGATCGCTCGGAATCGGAGCGCGATACCGTAAGACGAGGAGCGG 221
 DB 484 CGATAGAAGCGGATCGCTCGGAATCGGAGCGCGATACCGTAAGACGAGGAGCGG 425
 QY 222 TCAGCCCATTCGCGCGCAAGCTCTTCAGCAATATCAGGATAGCCACGCTATGTCTGA 281
 DB 424 TCAGCCCATTCGCGCGCAAGCTCTTCAGCAATATCAGGATAGCCACGCTATGTCTGA 365
 QY 282 TAGGGTCGCCACACCCAGCGCGGCACAGTCGATCAATCCAGAAAGCGGCATTTTC 341
 DB 364 TAGGGTCGCCACACCCAGCGCGGCACAGTCGATCAATCCAGAAAGCGGCATTTTC 305
 QY 342 ACCATGATATTCGCGAAGCAGGCGATCGCCATGGGTACGACAGAGATCCTCGCGTCGGGC 401
 DB 304 ACCATGATATTCGCGAAGCAGGCGATCGCCATGGGTACGACAGAGATCCTCGCGTCGGGC 245
 QY 402 ATCGCGCCCTTAGCCTGGGCAACAGTTGGCTGGCGCGAGCGCCCTGATGCTTTCGTCC 461
 DB 244 ATCGCGCCCTTAGCCTGGGCAACAGTTGGCTGGCGCGAGCGCCCTGATGCTTTCGTCC 185
 QY 462 AGATCATCTGATCGCAACAGCGGCTTCATCCGAGTACGTGCTCGCTCGATGCGATGT 521
 DB 184 AGATCATCTGATCGCAACAGCGGCTTCATCCGAGTACGTGCTCGCTCGATGCGATGT 125
 QY 522 TTCGTTGGTGGTGAATGGGCAAGGTAGCGGATCAAGCGTATGCAAGCGCGCGCATTCGA 581

Db 124 TTCGTTGGTGGTGAATGGGCAAGGTAGCGGATCAAGCGTATGAGCGCGCATTCGA 65
 QY 582 TCAGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 641
 Db 64 TCAGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5
 QY 642 GGCA 645
 Db 4 GGCA 1

RESULT 10
 LOCUS AQ875040
 DEFINITION V12054 mTn-3xHA/lacZ Insertion Library, strain Y278 Saccharomyces cerevisiae genomic 5', genomic survey sequence.
 ACCESSION AQ875040
 VERSION AQ875040.1 GI:6287284
 KEYWORDS GSS.
 SOURCE Saccharomyces cerevisiae (baker's yeast)
 ORGANISM Saccharomyces cerevisiae
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 REFERENCE 1 (bases 1 to 861)
 AUTHORS Ross-Macdonald, P., Roemer, T., Coelho, P. S. R., Agarwal, S., Kumar, A., deStages, S. A., Cheung, K. H., Sheehan, A., Symoniat, D., Jansen, R., Umansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Luge, R., Hager, K., Miller, P., Roeder, G. S., and Snyder, M.
 TITLE Large-Scale Analysis of the Yeast Genome by Transposon Tagging and Gene Disruption
 JOURNAL Unpublished
 COMMENT Contact: Kumar A
 Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
 Yale University
 P.O. Box 208103, New Haven, CT 06520-8103, USA
 Tel: 203 432 9949
 Fax: 203 432 6161
 Email: anuj.kumar@yale.edu
 te of mTn-3xHA/lacZ insertion.
 Seg primer: GGCCTTCTTCTTTTGAAGTAC
 Class: transposon-tagged.
 FEATURES Location/Qualifiers
 source 1..861
 /organism="Saccharomyces cerevisiae"
 /mol_type="genomic DNA"
 /strain="Y278 - S288C background, cir(0) rho(0)"
 /db_xref="taxon:4932"
 /lab_host="E. coli"
 /clone_lib="mTn-3xHA/lacZ Insertion Library, strain Y278"
 /note="Vector: pHSS6-Sal; A yeast genomic DNA library without 2 micron or mitochondrial DNA was prepared in pHSS6-Sal; Genomic DNA was size-fractionated (DNA of roughly 2-3 kb in length) prior to cloning. This library was subsequently mutagenized with a mTn-3xHA/lacZ minitransposon containing lacZ, URA3, and tet resistance."
 BASE COUNT 203 a 253 c 233 g 169 t 3 others
 ORIGIN

Query Match 49.3%; Score 493; DB 28; Length 861;
 Best Local Similarity 100.0%; Pred. No. 1.4e-247;
 Matches 493; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTGGAGGATCATCCAGCCGCGTCCCGAAACGATTCGGAAGCCCAACCTTTCATAG 60
 Db 215 CGCTGGAGGATCATCCAGCCGCGTCCCGAAACGATTCGGAAGCCCAACCTTTCATAG 274
 QY 61 AAGCGCGCGTGGATCGAATCTCGTATGCGAGTTGGCGCTCGCTTGGTGGTTCATT 120
 Db 275 AAGCGCGCGTGGATCGAATCTCGTATGCGAGTTGGCGCTCGCTTGGTGGTTCATT 334
 QY 121 TCGAACCCCGAGTCCCGCTCAGAAGAACTCGTCAAGAGCGCATAGAGCGCATGCCT 180

```

Db 335 TCAGACCCAGAGTCCCGCTCAGAGAACTCGTCAGAAAGGCGATAGAGGCGATCGCGT 394
QY 181 GCGAATCGGAGCGGCGATACCGTAAGACAGAGAGCGGTACGCCATTCGCGGCCAA 240
Db 395 GCGAATCGGAGCGGCGATACCGTAAGACAGAGAGCGGTACGCCATTCGCGGCCAA 454
QY 241 GCTCTTCAGCAATATCACGGGTAGCAACGGCTATGTCCTGTATAGCGGTCCGCCACCCA 300
Db 455 GCTCTTCAGCAATATCACGGGTAGCAACGGCTATGTCCTGTATAGCGGTCCGCCACCCA 514
QY 301 GCGGGCCACAGTCGATGATCCAGAAAGCGGCCATTTTCCACCATGATATTCGCGCAAGC 360
Db 515 GCGGGCCACAGTCGATGATCCAGAAAGCGGCCATTTTCCACCATGATATTCGCGCAAGC 574
QY 361 AGGCATCGGCATGCGGTACAGAGATCCTCGCGGTCCGGGATGCGGCTTGGACCTGG 420
Db 575 AGGCATCGGCATGCGGTACAGAGATCCTCGCGGTCCGGGATGCGGCTTGGACCTGG 634
QY 421 CGAACAGTTTCGGTGGCGGAGCCCTCGTATGCTCTTCGTCAGATCATCCTGATCGACAA 480
Db 635 CGAACAGTTTCGGTGGCGGAGCCCTCGTATGCTCTTCGTCAGATCATCCTGATCGACAA 694
QY 481 GACCGGCTTCCAT 493
Db 695 GACCGGCTTCCAT 707

```

```

RESULT 11
A0876220
LOCUS
DEFINITION
V152G7 mTn-3xHA/lacZ Insertion Library, strain Y2278 Saccharomyces
cerevisiae genomic 5', genomic survey sequence.
ACCESSION
A0876220
VERSION
A0876220.1 GI:6288464
KEYWORDS
GSS.
SOURCE
Saccharomyces cerevisiae (baker's yeast)
ORGANISM
Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
1 (bases 1 to 799)
Ross-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A.,
desEtages, S.A., Cheung, K.-H., Sheehan, A., Symoniat, D., Jansen, R.,
Umansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,
Hager, K., Miller, P., Roeder, G.S., and Snyder, M.
Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
Gene Disruption
Unpublished
Contact: Kumar A
Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
Yale University
P.O. Box 208103, New Haven, CT 06520-8103, USA
Tel: 203 432 9949
Fax: 203 432 6161
Email: anuj.kumar@yale.edu
te of mTn-3xHA/lacZ insertion.
Seq primer: GGCCTTCTTTTGGGAAGTAC
Class: transposon-tagged.
FEATURES
Location/Qualifiers
1..799
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/strain="Y2278 - S288C background, cir(0) rho(0)"
/db_xref="taxon:4932"
/lab_host="E. coli"
/clone_lib="mTn-3xHA/lacZ Insertion Library, strain Y2278"
/note="Vector: pHS6-Sal; A yeast genomic DNA library
without 2 micron or mitochondrial DNA was prepared in
pHS6-Sal; genomic DNA was size-fractionated (DNA of
roughly 2-3 kb in length) prior to cloning. This library
was subsequently mutagenized with a mTn-3xHA/lacZ
minitransposon containing lacZ, URA3, and tet resistance."

```

```

BASE COUNT 176 a 234 c 228 g 161 t
ORIGIN
Query Match 48.9%; Score 489; DB 28; Length 799;
Best Local Similarity 100.0%; Pred. No. 1.8e-245;
Matches 489; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGGAGGATCATCAGCGCGGTCCCGAAACGATTCCGAAAGCCACCTTTTCATAG 60
Db 187 CGTGGAGGATCATCAGCGCGGTCCCGAAACGATTCCGAAAGCCACCTTTTCATAG 246
QY 61 AAGCGCGGTGGAATCGAAATCTCGTGATGCGAGGTTGGCGGTTCGTTGGTCGTCATT 120
Db 247 AAGCGCGGTGGAATCGAAATCTCGTGATGCGAGGTTGGCGGTTCGTTGGTCGTCATT 306
QY 121 TCGAACCCACAGTCCCGCTCAGAGAACTCGTCAAGAAAGGCGATAGAGGCGATCGCGT 180
Db 307 TCGAACCCACAGTCCCGCTCAGAGAACTCGTCAAGAAAGGCGATAGAGGCGATCGCGT 366
QY 181 GCGAATCGGAGCGGCGATACCGTAAGACAGAGAGCGGTACGCCATTCGCGGCCAA 240
Db 367 GCGAATCGGAGCGGCGATACCGTAAGACAGAGAGCGGTACGCCATTCGCGGCCAA 426
QY 241 GCTCTTCAGCAATATCACGGGTAGCAACGGCTATGTCCTGTATAGCGGTCCGCCACCCA 300
Db 427 GCTCTTCAGCAATATCACGGGTAGCAACGGCTATGTCCTGTATAGCGGTCCGCCACCCA 486
QY 301 GCGGGCCACAGTCGATGATCCAGAAAGCGGCCATTTTCCACCATGATATTCGCGCAAGC 360
Db 487 GCGGGCCACAGTCGATGATCCAGAAAGCGGCCATTTTCCACCATGATATTCGCGCAAGC 546
QY 361 AGGCATCGGCATGCGGTACAGAGATCCTCGCGGTCCGGGATGCGGCTTGGACCTGG 420
Db 547 AGGCATCGGCATGCGGTACAGAGATCCTCGCGGTCCGGGATGCGGCTTGGACCTGG 606
QY 421 CGAACAGTTTCGGTGGCGGAGCCCTCGTATGCTCTTCGTCAGATCATCCTGATCGACAA 480
Db 607 CGAACAGTTTCGGTGGCGGAGCCCTCGTATGCTCTTCGTCAGATCATCCTGATCGACAA 666
QY 481 GACCGGCTT 489
Db 667 GACCGGCTT 675

```

```

RESULT 12
A0875907
LOCUS
DEFINITION
V130H3 mTn-3xHA/lacZ Insertion Library, strain Y2278 Saccharomyces
cerevisiae genomic 5', genomic survey sequence.
ACCESSION
A0875907
VERSION
A0875907.1 GI:6288151
KEYWORDS
GSS.
SOURCE
Saccharomyces cerevisiae (baker's yeast)
ORGANISM
Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
1 (bases 1 to 801)
Ross-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A.,
desEtages, S.A., Cheung, K.-H., Sheehan, A., Symoniat, D., Jansen, R.,
Umansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,
Hager, K., Miller, P., Roeder, G.S., and Snyder, M.
Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
Gene Disruption
Unpublished
Contact: Kumar A
Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
Yale University
P.O. Box 208103, New Haven, CT 06520-8103, USA
Tel: 203 432 9949
Fax: 203 432 6161
Email: anuj.kumar@yale.edu
te of mTn-3xHA/lacZ insertion.
Seq primer: GGCCTTCTTTTGGGAAGTAC

```

```

FEATURES
    source
        Class: transposon-tagged.
        Location/Qualifiers
            1..801
                /organism="Saccharomyces cerevisiae"
                /mol_type="genomic DNA"
                /strain="Y2278 - S288C background, cir(0) rho(0)"
                /db_xref="taxon:4932"
                /lab_host="E. coli"
                /clone_lib="mtn-3xHA/lacZ Insertion Library, strain Y2278"
                /note="vector: pHSSE-sal; A yeast genomic DNA library
                without 2 micron or mitochondrial DNA was prepared in
                pHSSE-sal; Genomic DNA was size-fractionated [DNA of
                roughly 2-3 kb in length] prior to cloning. This library
                was subsequently mutagenized with a mtn-3xHA/lacZ
                minitransposon containing lacZ, URA3, and tet resistance.

```

BASE COUNT	178 a	228 c	221 g	174 t	
Query Match	48.1%	Score 481	DB 28	Length 801	
Best Local Similarity	99.8%	Pred. No. 2.9e-241			
Matches 531; Conservative 0; Mismatches 1; Indels 0; Gaps 0;					
QY	1	CGCTGGAGGATCATCCAGCGCGGTCCCGAAACGATTCCGAAGCCCAACCTTTTCATAG	60		
Db	270	CGCTGGAGGATCATCCAGCGCGGTCCCGAAACGATTCCGAAGCCCAACCTTTTCATAG	329		
QY	61	AAGCGGCGGTGGAAATCGAAATCTCGTGATGGCAGGTTGGGGTGCCTTTGGTTCGGTCATT	120		
Db	330	AAGCGGCGGTGGAAATCGAAATCTCGTGATGGCAGGTTGGGGTGCCTTTGGTTCGGTCATT	389		
QY	121	TCGAACCCAGAGTCCCGCTCAGAGAACTCCTGTAAGAGGCGATAGAAGCGATGCGCT	180		
Db	330	TCGAACCCAGAGTCCCGCTCAGAGAACTCCTGTAAGAGGCGATAGAAGCGATGCGCT	449		
QY	181	GCGAATCGGAGCGGCATACCGTAAAGACACAGAGAGCGGTACGCCCATTTGCCGCGCAA	240		
Db	450	GCGAATCGGAGCGGCATACCGTAAAGACACAGAGAGCGGTACGCCCATTTGCCGCGCAA	509		
QY	241	GCTCTTCAGCNAATACAGGGTAGCCAACTGTTATGCTGTAGTCGGTTCGCCCAACCCA	300		
Db	510	GCTCTTCAGCNAATACAGGGTAGCCAACTGTTATGCTGTAGTCGGTTCGCCCAACCCA	569		
QY	301	CGCGGCCACAGTCGATGAATCCAGAAAAGCGGCCATTTCCACCATGATATTGGCAAGC	360		
Db	570	CGCGGCCACAGTCGATGAATCCAGAAAAGCGGCCATTTCCACCATGATATTGGCAAGC	629		
QY	361	AGGCATCGCCATGGGTCAAGCAGAGATCTTCGCCGTGGGCAATGCGCGCTTGAGCTGG	420		
Db	630	AGGCATCGCCATGGGTCAAGCAGAGATCTTCGCCGTGGGCAATGCGCGCTTGAGCTGG	689		
QY	421	CGAACAGTTCGGTGGCGAGCGCCCTGATGCTCTTCGTCCAGAGATCATCTGTATCGACAA	480		
Db	690	CGAACAGTTCGGTGGCGAGCGCCCTGATGCTCTTCGTCCAGAGATCATCTGTATCGACAA	749		
QY	481	GACGGCTTCATCCGAGTACGTGCTCGCTCGATCGGATGTTTCGTTGGTG	532		
Db	750	GACGGCTTCATCCGAGTACGTGCTCGCTCGATCGGATGTTTCGTTGGTG	801		

RESULT 13	
AQ876024	
LOCUS	783 bp DNA linear GSS 08-NOV-1999
DEFINITION	V132C9 mTn-3xHA/lacZ Insertion Library, strain Y2278 Saccharomyces cerevisiae Genomic 5', genomic survey sequence.
ACCESSION	AQ876024
VERSION	AQ876024.1
KEYWORDS	GI:6288268
SOURCE	GSS.
ORGANISM	Saccharomyces cerevisiae (baker's yeast)
	Saccharomyces cerevisiae
	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
REFERENCE	1 (bases 1 to 783)

AUTHORS	TITLE	JOURNAL	COMMENT
...

Ross-Macdonald, P., Roemer, T., Coelho, P. S. R., Agarwal, S., Kumar, A.,
destages, S. A., Cheung, K.-H., Sheehan, A., Symoniatis, D., Jansen, R.,
Umsky, L., Heidman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,
Hager, K., Miller, P., Roeder, G. S. and Snyder, M.
Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
Gene Disruption
Unpublished
Contact: Kumar A
Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
Yale University
P.O. Box 208103, New Haven, CT 06520-8103, USA
Tel: 203 432 9949
Fax: 203 432 6161
Email: anuj.kumar@yale.edu
te of mtn-3xHA/lacZ insertion.

FEATURES
source

```

1: 705
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/strain="Y2278 - S288C background, cir(0) rho(0)"
/db_xref="taxon:4932"
/lab_host="E. coli"
/clone_lib="mrn-3xHA/lacZ Insertion Library, strain Y2278"
/notes="Vector: PHSS6-Sal; A yeast genomic DNA library without 2 micron or mitochondrial DNA was prepared in PHSS6-Sal; genomic DNA was size-fractionated (DNA of roughly 2-3 kb in length) prior to cloning. This library was subsequently mutagenized with a mrn-3xHA/lacZ minitransposon containing lacZ, URA3, and tet resistance."
181 a 224 c 223 g 154 t 1 others
BASE COUNT
ORIGIN

```

Query Match	47.9%;	Score 479;	DB 28;	Length 783;
Best Local Similarity	100.0%;	Pred. No. 3.2e-240;		
Matches 479;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CGCTGAGGATCATTCAGCGCGGTCCGGAAACGATTCGGAAGCCCAACCTTTCATAG	60	
Db	296	CGTGGAGATCATTCAGCGCGGTCCGGAAACGATTCGGAAGCCCAACCTTTCATAG	355	
QY	61	AAGCGCGGTGGAATCGAAATCTGTGATGGCAGGTTGGCGTTCGTCGGTCATT	121	
Db	356	AAGCGCGGTGGAATCGAAATCTGTGATGGCAGGTTGGCGTTCGTCGGTCATT	415	
QY	121	TGGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGCGCATAGAGGCGCATCGCT	180	
Db	416	TGGAACCCAGAGTCCCGCTCAGAGAACTCGTCAAGAGGCGCATAGAGGCGCATCGCT	475	
QY	181	GGGAATCGGAGCGGCGATACCGTAAAGCAGAGGAGCGGTTCAGCCCATTCGCCGCCAA	240	
Db	476	GGGAATCGGAGCGGCGATACCGTAAAGCAGAGGAGCGGTTCAGCCCATTCGCCGCCAA	535	
QY	241	GCTCTTCAGCAATATCACGGGTAGCCAAACGCTATGTCCTGATAGCGGTTCGGCACACCCA	300	
Db	536	GCTCTTCAGCAATATCACGGGTAGCCAAACGCTATGTCCTGATAGCGGTTCGGCACACCCA	595	
QY	301	GCGGGCCAAGTCGATGAATCCAGAAAACGGGCCATTTTCCACCATGATATTCGGAAGC	360	
Db	596	GCGGGCCAAGTCGATGAATCCAGAAAACGGGCCATTTTCCACCATGATATTCGGAAGC	655	
QY	361	AGGCATCGGCATGGGTTCAGACGAGATCCTCGCGGTTCGGGCATCGCGCTTGAGCGCTGG	420	
Db	656	AGGCATCGGCATGGGTTCAGACGAGATCCTCGCGGTTCGGGCATCGCGCTTGAGCGCTGG	715	
QY	421	CGAAAGTTTCGGCTCGCGCGAGCCCTGATGCTCTTCGTCCAGATCATCTGATGACA	479	
Db	716	CGAAAGTTTCGGCTCGCGCGAGCCCTGATGCTCTTCGTCCAGATCATCTGATGACA	774	

RESULT 14

AQ447874/c
 LOCUS 561 bp DNA linear GSS 08-APR-1999
 DEFINITION clone mgxb0012101f CUGI Rice Blast BAC Library Magnaporthe grisea genomic
 clone mgxb0012101f, genomic survey sequence.
 ACCESSION AQ447874
 VERSION 1
 KEYWORDS GSS
 SOURCE Magnaporthe grisea (anamorph: Pyricularia grisea)
 ORGANISM Magnaporthe grisea
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
 1 (bases 1 to 561)
 Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
 Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
 A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
 Genome
 JOURNAL Unpublished
 COMMENT Contact: Dean RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson University, Clemson, SC 29634
 Tel: 864 656 5737
 Fax: 864 656 4293
 Email: rdean@clemson.edu
 Seq primer: TAATACGACTCACTATAGGG
 Class: BAC ends
 High quality sequence start: 42
 High quality sequence stop: 326.
 High quality sequence stop: 326.
 Location/Qualifiers
 1..561
 /organism="Magnaporthe grisea"
 /mol_type="genomic DNA"
 /strain="70-15"
 /db_xref="taxon:148305"
 /clone="mgxb0012101f"
 /tissue_type="Protoplasts"
 /lab_host="E. coli DH10B"
 /clone_lib="CUGI Rice Blast BAC Library"
 /note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
 Rice blast is one of the most devastating fungal diseases
 of rice world wide. It is a filamentous ascomycete with
 a haploid genome (n=7) of approximately 40 Mbp. Rice
 blast is an important model fungal pathogen for studying
 numerous aspects of the fungal-host interaction. In
 order to facilitate genome wide analysis, a BAC library
 containing 9216 clones with an average insert size of 130
 kbp was constructed. This library represents greater
 than 25X genome coverage. High density colony filters
 are available upon request."
 104 a 163 c 174 g 118 t 2 others
 BASE COUNT 104 a 163 c 174 g 118 t 2 others
 ORIGIN
 Query Match 46.8%; Score 468; DB 28; Length 561;
 Best Local Similarity 100.0%; Pred. No. 1.8e-234;
 Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 294 ACACCCAGCGGCGACAGTCGATGATTCAGAAAGCGGCATTTCCACCATGATTC 353
 Db 561 ACACCCAGCGGCGACAGTCGATGATTCAGAAAGCGGCATTTCCACCATGATTC 502
 Qy 354 GGCAAGCAGGCATCGCATGGTGCAGCAGATCTTCGGCGCGGCATGCGCGCTTG 413
 Db 501 GGCAGCAGGCATCGCATGGTGCAGCAGATCTTCGGCGCGGCATGCGCGCTTG 442
 Qy 414 AGCCTGCGAACAGTTCGGTGGCGGAGCCCTTGATGCTTCGTCAGATCATCTCGA 473
 Db 441 AGCCTGCGAACAGTTCGGTGGCGGAGCCCTTGATGCTTCGTCAGATCATCTCGA 382
 Qy 474 TCACACAGCGGCTTCATCCAGTACGTCGCTCGATCGATGTTTCGTTGGTGG 533
 Db 381 TCACACAGCGGCTTCATCCAGTACGTCGCTCGATCGATGTTTCGTTGGTGG 322
 Qy 534 TCGAATGGCAGGTAGCCGGATCAAGCGTATGCGAGCGCGCATTCGATCAGCCATGATG 593

321 TCGAATGGCAGGTAGCCGGATCAAGCGTATGCGCGCCCGCATTCGATCAGCCATGATG 262
 Qy 594 GATCTTTCTCGGAGGAGCAAGTGTGATGACACAGGAGATCCTGCCGCGCACTTCGCCC 653
 Db 261 GATCTTTCTCGGAGGAGCAAGTGTGATGACACAGGAGATCCTGCCGCGCACTTCGCCC 202
 Qy 654 AATAGCAGCCAGTCCCTTCCCGCTTCAGTTCAGTACAAACGTCGAGCAGCTGGCCAAAGGAACG 713
 Db 201 AATAGCAGCCAGTCCCTTCCCGCTTCAGTTCAGTACAAACGTCGAGCAGCTGGCCAAAGGAACG 142
 Qy 714 CCGCTCGTGGCAGCAGCAGATAGCGCGCTGCCTCTGCTCGCTTCAGTTCA 761
 Db 141 CCGCTCGTGGCAGCAGCAGATAGCGCGCTGCCTCTGCTCGCTTCAGTTCA 94
 RESULT 15
 AQ398387/c
 LOCUS 620 bp DNA linear GSS 06-MAR-1999
 DEFINITION clone mgxb0009005f CUGI Rice Blast BAC Library Magnaporthe grisea genomic
 clone mgxb0009005f, genomic survey sequence.
 ACCESSION AQ398387
 VERSION 1
 KEYWORDS GSS
 SOURCE Magnaporthe grisea (anamorph: Pyricularia grisea)
 ORGANISM Magnaporthe grisea
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
 1 (bases 1 to 620)
 Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
 Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
 A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
 Genome
 JOURNAL Unpublished
 COMMENT Contact: Dean RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson University, Clemson, SC 29634
 Tel: 864 656 5737
 Fax: 864 656 4293
 Email: rdean@clemson.edu
 Seq primer: TAATACGACTCACTATAGGG
 Class: BAC ends
 High quality sequence stop: 187.
 High quality sequence stop: 187.
 Location/Qualifiers
 1..620
 /organism="Magnaporthe grisea"
 /mol_type="genomic DNA"
 /strain="70-15"
 /db_xref="taxon:148305"
 /clone="mgxb0009005f"
 /tissue_type="Protoplasts"
 /lab_host="E. coli DH10B"
 /clone_lib="CUGI Rice Blast BAC Library"
 /note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
 Rice blast is one of the most devastating fungal diseases
 of rice world wide. It is a filamentous ascomycete with
 a haploid genome (n=7) of approximately 40 Mbp. Rice
 blast is an important model fungal pathogen for studying
 numerous aspects of the fungal-host interaction. In
 order to facilitate genome wide analysis, a BAC library
 containing 9216 clones with an average insert size of 130
 kbp was constructed. This library represents greater
 than 25X genome coverage. High density colony filters
 are available upon request."
 121 a 172 c 188 g 139 t
 BASE COUNT 121 a 172 c 188 g 139 t
 ORIGIN
 Query Match 46.8%; Score 468; DB 28; Length 620;
 Best Local Similarity 100.0%; Pred. No. 1.9e-234;
 Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 245 TTCAGCAATATCAGCGGTAGCCCAACGATGATCTCTGATGATCGGTCCGCCACACCCAGCG 304

Db	468	TTTCAGCAATATCAGGGTAGCCACAGCTATGTCTGTATAGGGTCCGCCACACCCAGCCG	409
Qy	305	GCACAGTCGATGATCCAGAAAGCGCCATTTCCACCATGATATTGGCAAGCAGGC	364
Db	408	GCACAGTCGATGATCCAGAAAGCGCCATTTCCACCATGATATTGGCAAGCAGGC	349
Qy	365	ATCGCCATGGGTCAAGACGAGATCCTCGCCGTCCGGCATGCGGCTTGAGCCTGGCGAA	424
Db	348	ATCGCCATGGGTCAAGACGAGATCCTCGCCGTCCGGCATGCGGCTTGAGCCTGGCGAA	289
Qy	425	CAGTTCCGCTGGCGAGCCCTGATGCTTTTCGTCAGATCATCTGATCGACAAGACC	484
Db	288	CAGTTCCGCTGGCGAGCCCTGATGCTTTTCGTCAGATCATCTGATCGACAAGACC	229
Qy	485	GGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTTCGTTGGTGGTGGCA	544
Db	228	GGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTTCGTTGGTGGTGGCA	169
Qy	545	GGTAGCCGATCAAGCGTATGACGCGCGCATTCGATCAGCCATGATGGATATTTCTC	604
Db	168	GGTAGCCGATCAAGCGTATGACGCGCGCATTCGATCAGCCATGATGGATATTTCTC	109
Qy	605	GGCAGGAGCAAGTGATGACAGAGATCCTGCGCCCGCATTCGCCCAATAGCAGCCA	664
Db	108	GGCAGGAGCAAGTGATGACAGAGATCCTGCGCCCGCATTCGCCCAATAGCAGCCA	49
Qy	665	GTCCCTTCCCGCTTCAGTGACAAACGTTCGAGCACAGCTGGCAAGAAC	712
Db	48	GTCCCTTCCCGCTTCAGTGACAAACGTTCGAGCACAGCTGGCAAGAAC	1

Search completed: February 16, 2004, 03:15:50
Job time : 2641.53 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: February 25, 2004, 16:56:30 ; Search time 6417 Seconds
(without alignments)
11306.878 Million cell updates/sec

Title: US-09-921-143-1
Perfect score: 1674
Sequence: 1 gtcctccaccatgactgctg.....ttataaaaaaaaaaaaaaa 1674

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.man.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1674	100.0	1674	6	BD270607	BD270607 Vascular
2	1674	100.0	1674	6	AR381510	AR381510 Sequence
3	1674	100.0	1674	6	BD001745	BD001745 Human vas
4	1674	100.0	1674	6	BD134149	BD134149 Vascular
5	1666	99.5	1948	9	BC035212	BC035212 Homo sapi
6	1664.6	99.4	2031	6	AR231261	AR231261 Sequence
7	1664.6	99.4	2031	6	AR342826	AR342826 Sequence
8	1664.6	99.4	2031	6	AR342826	AR342826 Sequence
9	1663	99.3	2031	6	AR231262	AR231262 Sequence
10	1663	99.3	2031	6	AR342827	AR342827 Sequence
11	1651.8	98.7	1997	6	AR112510	AR112510 Sequence
12	1651.8	98.7	1997	6	AR157589	AR157589 Sequence
13	1651.8	98.7	1997	6	BD247178	BD247178 Use of VE
14	1651.8	98.7	1997	6	AR201986	AR201986 Sequence
15	1651.8	98.7	1997	6	AR212635	AR212635 Sequence
16	1651.8	98.7	1997	6	AR430033	AR430033 Sequence
17	1651.8	98.7	1997	6	AX234371	AX234371 Sequence
18	1651.8	98.7	1997	6	AX743106	AX743106 Sequence
19	1651.8	98.7	1997	6	BD082159	BD082159 Vascular
20	1651.8	98.7	1997	6	HSVEGFC	X94216 H.sapiens m
21	1651.8	98.7	2015	9	HSU43142	U43142 Human vascu
22	1651.8	98.7	2679	6	BD247192	BD247192 Use of VE
23	1620.6	96.8	1939	9	HSU58111	U58111 Human FLT4
24	1526	91.2	1526	6	BD270608	BD270608 Vascular
25	1526	91.2	1526	6	BD134150	BD134150 Vascular
26	1292.6	77.2	1777	4	AB004275	AB004275 Bos tauru
27	1256.8	75.1	1260	6	AX481509	AX481509 Sequence
28	1212.4	72.4	1804	10	MMU58112	U73620 Mus musculu
29	1182.2	70.6	1818	10	MMU73620	U73620 Mus musculu
30	1182.2	70.6	1836	6	AR112511	AR112511 Sequence
31	1182.2	70.6	1836	6	AR201987	AR201987 Sequence
32	1182.2	70.6	1836	6	AR212641	AR212641 Sequence
33	1182.2	70.6	1836	6	AR430039	AR430039 Sequence
34	1182.2	70.6	1836	6	BD082160	BD082160 Vascular
35	1138.4	68.0	1140	6	AR157586	AR157586 Sequence
36	1076.4	64.3	1596	10	AY032729	AY032729 Rattus no
37	860.8	51.4	1218	10	AF432867	AF432867 Meriones
38	805.6	48.1	1257	5	CCV15937	Y15937 Coturnix co
39	804.2	48.0	1741	6	AR112512	AR112512 Sequence
40	804.2	48.0	1741	6	AR201988	AR201988 Sequence
41	804.2	48.0	1741	6	BD082161	BD082161 Vascular
42	486.4	29.1	126929	9	AC093801	AC093801 Homo sapi
43	381	22.8	1936	5	AF466147	AF466147 Danilo rer
44	335.6	20.0	177771	2	AC140905	AC140905 Homo sapi
45	321	19.2	378	4	AF009178	AF009178 Bos tauru

ALIGNMENTS

RESULT 1	BD270607	1674 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD270607	Vascular endothelial growth factor 2.			
DEFINITION	BD270607				
ACCESSION	BD270607.1	GI:33080375			
VERSION	JP 2002539082-A/1.				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 1674)				
AUTHORS	Rosen, C.A., Alderson, R., Melder, R., Roschke, V. and Ruben, S.M.				
TITLE	Vascular endothelial growth factor 2				
JOURNAL	Patent: JP 2002539082-A 1 19-NOV-2002;				

COMMENT		HUMAN GENOME SCIENCES INC		OS Homo sapiens (human)		PN JP 2002539082-A/1		PD 13-NOV-2002		PF 07-FEB-2000 JP 2000596954		PR 08-FEB-1999 US 60/119179,12-FEB-1999 US 60/119926 PR		03-JUN-1999 US 60/137796,22-DEC-1999 US 60/171505 PI CRAIG		A ROSEN, RALPH ALDERSON, ROBERT MELDER, VIKTOR ROSCHKE, PI STEVEN M RUBEN	
PC A61K38/22,A61K38/00,A61K38/21,A61K38/27,A61K39/395,A61K39/395, A61K45/00,A61K47/48,A61K48/00,A61P3/10,A61P25/02,A61P27/02, PC A61P29/00, PC A61P31/12,A61P37/00,C12N5/10//A61K35/12,C07K16/18,C12N15/09, PC A61K37/24, PC C12N5/00,A61K37/02,A61K37/66,A61K37/36,C12N15/00 CC Vascular endothelial growth factor 2		FH Key		Location/Qualifiers		FT CDS		(12)...(1268)		FT sig_peptide		(12)...(80)		FT mat_peptide		(81)...(1268)	
FEATURES		source		1..1674		/organism="Homo sapiens"		/mol_type="genomic DNA"		/db_xref="taxon:9606"							
ORIGIN		Query March		100.0%; Score 1674; DB 6; Length 1674;		Best Local Similarity		100.0%; Pred No. 0;		Matches 1674; Conservative		0; Mismatches		0; Indels		0; Gaps	
QY	1	GTCTTCCACCATGCACTGCTGGCTTCTTCTGTCGCTGTTCTCTGCTGCGCTG 60															
DB	1	GTCTTCCACCATGCACTGCTGGCTTCTTCTGTCGCTGTTCTCTGCTGCGCTG 60															
QY	61	CGCTGCTCCGGTCTCCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120															
DB	61	CGCTGCTCCGGTCTCCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120															
QY	121	ACCTCTCGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180															
DB	121	ACCTCTCGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180															
QY	181	AGGAGCAGTACGGTCTGTGTCAGTGTAGAGTAAGTAAGTAAGTAAGTAAG 240															
DB	181	AGGAGCAGTACGGTCTGTGTCAGTGTAGAGTAAGTAAGTAAGTAAGTAAG 240															
QY	241	ATTGGAATGTACAAAGTGTACAGTAAAGAAAGAGGCTGGCAACATACAGAG 300															
DB	241	ATTGGAATGTACAAAGTGTACAGTAAAGAAAGAGGCTGGCAACATACAGAG 300															
QY	301	CCAACTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGCAGCATTATA 360															
DB	301	CCAACTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGCAGCATTATA 360															
QY	361	AGATCTTGAAGATATTGATAATAGTGGAGAGAGACTCAATGCGAGGAGTGT 420															
DB	361	AGATCTTGAAGATATTGATAATAGTGGAGAGAGACTCAATGCGAGGAGTGT 420															
QY	421	GTATAGATGGGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 480															
DB	421	GTATAGATGGGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 480															
QY	481	TGTCGCTACAGAGTGGGCTTCTGCAATAGTAGAGGAGTGGAGTGGAGTGGAG 540															
DB	481	TGTCGCTACAGAGTGGGCTTCTGCAATAGTAGAGGAGTGGAGTGGAGTGGAG 540															
QY	541	GCACGAGTACCTCAGCAGAGAGTATTATTGAATTTACAGTGGCTCTCTCAAGGCCCA 600															
DB	541	GCACGAGTACCTCAGCAGAGAGTATTATTGAATTTACAGTGGCTCTCTCAAGGCCCA 600															
QY	601	ACCAGTACATCAGTTTGGCAATCAGACTTCTCTCGCGATGCAATGCTAAACTGGATG 660															

RESULT 2

AR381510
LOCUS AR381510 1674 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1 from patent US 6608182.
ACCESSION AR381510
VERSION AR381510.1 GI:40089598
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1674)
AUTHORS Rosen, C., Hu, J.-S. and Cao, L.
TITLE Human vascular endothelial growth factor 2
JOURNAL Patent: US 6608182-A 1 19-AUG-2003;
FEATURES
Location/Qualifiers
1..1674
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 1674; DB 6; Length 1674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GTCCTCCACCATGCACTCGCTGGGCTTCTCTCTGTGGGGTGTCTCTCTCTCGCGCTG 60
Db 1 GTCCTCCACCATGCACTCGCTGGGCTTCTCTCTGTGGGGTGTCTCTCTCTCGCGCTG 60
Qy 61 CGTGTCTCCCGGTCCTCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 120
Db 61 CGTGTCTCCCGGTCCTCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 120
Qy 121 ACCTCTCGGCGCGGCG 180
Db 121 ACCTCTCGGCGCGGCG 180
Qy 181 AGGAGCAGTTACGGTCTGTGTCCAGTGTAGATGAATCATGACTGTACTTACCCAGAT 240
Db 181 AGGAGCAGTTACGGTCTGTGTCCAGTGTAGATGAATCATGACTGTACTTACCCAGAT 240
Qy 241 ATTGGAAATGTACAGTGTCTAGTCTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Db 241 ATTGGAAATGTACAGTGTCTAGTCTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Qy 301 CCAACCTCAACTCAAGGAG 360
Db 301 CCAACCTCAACTCAAGGAG 360
Qy 361 AGATCTTGAAAGATTGATTAATGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 361 AGATCTTGAAAGATTGATTAATGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Qy 421 GTATAGATGTGGGAG 480
Db 421 GTATAGATGTGGGAG 480
Qy 481 TGTCCGCTCAGAGATGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db 481 TGTCCGCTCAGAGATGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Qy 541 GCAGAGCTACTCAGCAG 600
Db 541 GCAGAGCTACTCAGCAG 600
Qy 601 AACCAAGTAACTAGTTTGGCAATCACACTTCTCTGCGGATGATGCTAAACTGGATG 660
Db 601 AACCAAGTAACTAGTTTGGCAATCACACTTCTCTGCGGATGATGCTAAACTGGATG 660
Qy 661 TTATCAGACAGTTCAATTCATATTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 661 TTATCAGACAGTTCAATTCATATTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Qy 721 AGGAGCGGAG 780
Db 721 AGGAGCGGAG 780

Db 721 AGGAGCGGAG 780
Qy 781 TGGCTCAGAGAGATTTTATGTTTCTCGAGTCTGGAGATGACTCAACAGAGATGATTC 840
Db 781 TGGCTCAGAGAGATTTTATGTTTCTCGAGTCTGGAGATGACTCAACAGAGATGATTC 840
Qy 841 ATGACATCTGTGGACCAACAGAGAGCTGGATGAGAGACCTGTCAGTGTCTCTCAGAG 900
Db 841 ATGACATCTGTGGACCAACAGAGAGCTGGATGAGAGACCTGTCAGTGTCTCTCAGAG 900
Qy 901 CGGGCTTTCGGCTCGCAGCTGTGGACCCCAAGAACTAGACAGAACTCATCCAGT 960
Db 901 CGGGCTTTCGGCTCGCAGCTGTGGACCCCAAGAACTAGACAGAACTCATCCAGT 960
Qy 961 GTGTCTGTAATAAAACAACTCTTCCCGAGCCCAATGTGGGGCCCAACGAGAAATTTGATGAA 1020
Db 961 GTGTCTGTAATAAAACAACTCTTCCCGAGCCCAATGTGGGGCCCAACGAGAAATTTGATGAA 1020
Qy 1021 ACACATGCCAGTGTGTATGTAATAAGAACTGCCCCAGAAATCAACCCCTAAATCTCGGAA 1080
Db 1021 ACACATGCCAGTGTGTATGTAATAAGAACTGCCCCAGAAATCAACCCCTAAATCTCGGAA 1080
Qy 1081 AATGTCCCTGTGAATGTACAGAACTCCACAGAAATGCTTGTAAAGGAAAGAGATTC 1140
Db 1081 AATGTCCCTGTGAATGTACAGAACTCCACAGAAATGCTTGTAAAGGAAAGAGATTC 1140
Qy 1141 ACCACCAACATGCGAGCTGTACAGCGGCCATGTACGAAACCGCCAGAGAGCTTGTGAGC 1200
Db 1141 ACCACCAACATGCGAGCTGTACAGCGGCCATGTACGAAACCGCCAGAGAGCTTGTGAGC 1200
Qy 1201 CAGGATTTTCATATAGTGAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
Db 1201 CAGGATTTTCATATAGTGAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
Qy 1261 AATGAGCTAGATGTACTGT 1320
Db 1261 AATGAGCTAGATGTACTGT 1320
Qy 1321 TGCCACAGTAGAATCTGTGTGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 1321 TGCCACAGTAGAATCTGTGTGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Qy 1381 AAAGTCTGTCTTCTGTAACCATGT 1440
Db 1381 AAAGTCTGTCTTCTGTAACCATGT 1440
Qy 1441 CAAAAGGCTCTGTGTGAAG 1500
Db 1441 CAAAAGGCTCTGTGTGAAG 1500
Qy 1501 TTGTGATTTCTTTTAAAGAACTGAT 1560
Db 1501 TTGTGATTTCTTTTAAAGAACTGAT 1560
Qy 1561 ATTCATTTTATAGCAACCAATTTGGTAAACTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1620
Db 1561 ATTCATTTTATAGCAACCAATTTGGTAAACTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1620
Qy 1621 GCAATAATGTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1674
Db 1621 GCAATAATGTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1674

RESULT 3
BD001745
LOCUS Human vascular endothelial growth factor 2.
DEFINITION BD001745
ACCESSION BD001745.1 GI:18626304
VERSION JP 2000069982-A/1.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

LOCUS	BD134149	1674 bp	DNA	linear	PAT 18-SEP-2002
DEFINITION	Vascular endothelial growth factor 2.				
ACCESSION	BD134149				
VERSION	BD134149.1	GI:232229094			
KEYWORDS	JP 2002505873-A/1.				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 1674)				
AUTHORS	Rosen,C.A., Kao,L. and Fu,J.S.				
TITLE	Vascular endothelial growth factor 2				
JOURNAL	Patent: JP 2002505873-A 1 26-FEB-2002;				
COMMENT	HUMAN GENOME SCIENCES INC				
	OS Unidentified				
	PN JP 2002505873-A/1				
	PF 26-FEB-2002				
	PP 10-MAR-1999 JP 2000535731				
	PR 13-MAR-1998 US 09/042105,30-JUN-1998 US 09/107997 PI				
	CRAIG A ROSEN,LIANG KAO, JTN SHAN FU				
	PC C12N1/21,A61K38/22,A61K48/00,A61P9/10,A61P9/12,A61P17/06, PC				
	A61P19/02				
	PC C07K14/50,C07K16/22,C12N5/10,C12N15/09,C12Q1/68,C12Q1/70, PC				
	A61K37/24,				
	PC C12N5/00,C12N15/00				
	CC Strandedness: Double;				
	CC Topology: Linear;				
	CC Vascular endothelial growth factor 2				
	PH Key location/Qualifiers				
	FT sig_peptide 12..80				
	FT mat_peptide 91..1268				
	FT CDS 12..1268.				
FEATURES	Location/Qualifiers				
source	1..1674				
	/organism="unidentified"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:32644"				
ORIGIN					
Query Match	100.0%;	Score 1674;	DB 6;	Length 1674;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 1674;	Conservative	0;	Mismatches	0;	Gaps 0;
QY	1	GTCTTCCACATGCACTCGCTGGGCTTCTTCTCTGTGGCGTGTCTCTGTCTCGCGCTG 60			
Db	1	GTCTTTCACATGCACTCGCTGGGCTTCTCTCTGTGGCGTGTCTCTGTCTCGCGCTG 60			
QY	61	CGCTGTCTCCCGGTCTCTCGAGGCGCCGCGCGCGCGCTTCGAGTCCGGAATCG 120			
Db	61	CGTGTCTCCCGGTCTCTCGAGGCGCCGCGCGCGCTTCGAGTCCGGAATCG 120			
QY	121	ACCTCTCGAGCGGAGCCGAGCCGCGGCGAGCCACGGCTTATGCAAGCAAGATCTGG 180			
Db	121	ACCTCTCGAGCGGAGCCGCGGCGAGCCACGGCTTATGCAAGCAAGATCTGG 180			
QY	181	AGAGAGAGTTTACGGTCTGTGTCCAGTGTAGATGAATCATGACTGTACTCTACCCAGAT 240			
Db	181	AGAGAGAGTTTACGGTCTGTGTCCAGTGTAGATGAATCATGACTGTACTCTACCCAGAT 240			
QY	241	ATTGGAATATGTACAAAGTGTACAGTAAAGAAAGAGGCTGGCAACATAACAGAGAACAG 300			
Db	241	ATTGGAATATGTACAAAGTGTACAGTAAAGAAAGAGGCTGGCAACATAACAGAGAACAG 300			
QY	301	CCAACTCAACTCAAGGACGAGAGACTATAAAATTGCTGAGCACATTTATATACAG 360			
Db	301	CCAACTCAACTCAAGGACGAGAGACTATAAAATTGCTGAGCACATTTATATACAG 360			
QY	361	AGATCTTTGAAAAGTATTGATATGATGATGAGTGGAGAAAGACTCAATGTCATGCCAGGGAGGTG 420			
Db	361	AGATCTTTGAAAAGTATTGATATGATGATGAGTGGAGAAAGACTCAATGTCATGCCAGGGAGGTG 420			
QY	421	GATATGATGTGGGAGAGGAGTTTGGAGTCCGCAAAACACCTCTCTTTAACTCCATGTG 480			
Db	421	GATATGATGTGGGAGAGGAGTTTGGAGTCCGCAAAACACCTCTCTTTAACTCCATGTG 480			

Db	735	TGTCGCTCTACAGATGTGGGGGTGCTGCAATAGTATGAGGGGCTGCAGTGCATGATCAACCA	794
Qy	541	GCACGAGCTACCTCAGCAAGACGTTATTATTGAAATTTACAGTGCCTCTCTCTCAAGGCCCA	600
Db	795	GCACGAGCTACCTCAGCAAGACGTTATTATTGAAATTTACAGTGCCTCTCTCTCAAGGCCCA	854
Qy	601	AACCAAGTAACAATCAGTTTGGCAATCAACATTCCTGCGGATGGAATGCTAAATCTGGATG	660
Db	855	AACCAAGTAACAATCAGTTTGGCAATCAACATTCCTGCGGATGGAATGCTAAATCTGGATG	914
Qy	661	TTTACAGACAAGTTCATTTCCTAATTAGACGTTCCCTGCCAGCAACACTACACAGTGTCT	720
Db	915	TTTACAGACAAGTTCATTTCCTAATTAGACGTTCCCTGCCAGCAACACTACACAGTGTCT	974
Qy	721	AGGCAGCAACAAGACCTGCGCCCAACAAATTACATGTGGAATAATCACTCTGCAGATGCC	780
Db	975	AGGCAGCAACAAGACCTGCGCCCAACAAATTACATGTGGAATAATCACTCTGCAGATGCC	1034
Qy	781	TGGCTCAGAAAGATTTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGGATTCC	840
Db	1035	TGGCTCAGAAAGATTTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGGATTCC	1094
Qy	841	ATGACATCTGTGGACAAACAAGGAGCTGGATGAGAGACCTGTCAGTGTGCTCTGCAGAG	900
Db	1095	ATGACATCTGTGGACAAACAAGGAGCTGGATGAGAGACCTGTCAGTGTGCTCTGCAGAG	1154
Qy	901	CGGGCTTCGGCTTCGCCAGCTGTGGACCCCAACAAGAACTAGACAGAAACTCATGCCAGT	960
Db	1155	CGGGCTTCGGCTTCGCCAGCTGTGGACCCCAACAAGAACTAGACAGAAACTCATGCCAGT	1214
Qy	961	GTGCTCTGTAACAAAACAACTCTTCCGCCAGCCAAATGTGGGGCCAAACCGAGAAATTTGATCAA	1020
Db	1215	GTGCTCTGTAACAAAACAACTCTTCCGCCAGCCAAATGTGGGGCCAAACCGAGAAATTTGATCAA	1274
Qy	1021	ACACATGCCAGTGTATGTATGTAAGAAACCTGCCCCAGAAATCAACCCCTAAATCTCGGAA	1080
Db	1275	ACACATGCCAGTGTATGTATGTAAGAAACCTGCCCCAGAAATCAACCCCTAAATCTCGGAA	1334
Qy	1081	AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTGTTTAAAGGAAAGAAAGTTC	1140
Db	1335	AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTGTTTAAAGGAAAGAAAGTTC	1394
Qy	1141	ACCACCAACAATGCAGCTGTTTACAGCGGCCATGTACGAAACCGCCAGAAAGCTGTGAGC	1200
Db	1395	ACCACCAACAATGCAGCTGTTTACAGCGGCCATGTACGAAACCGCCAGAAAGCTGTGAGC	1454
Qy	1201	CAGGATTTTCATATAGTGAAGAGTGTGCTGTGTGTCCTTCATATGTCGAAGACCA	1260
Db	1455	CAGGATTTTCATATAGTGAAGAGTGTGCTGTGTGTCCTTCATATGTCGAAGACCA	1514
Qy	1261	AAATGAGCTAAGATGTACTGTTTCCAGTTTCATCGATTTTCTATTATGAAACCTGTGT	1320
Db	1515	AAATGAGCTAAGATGTACTGTTTCCAGTTTCATCGATTTTCTATTATGAAACCTGTGT	1574
Qy	1321	TGCCACAGTAGAAGCTGCTGTGACAGAGAGCCCTGTGGTCCATGCTGCTAAACAAGACA	1380
Db	1575	TGCCACAGTAGAAGCTGCTGTGACAGAGAGCCCTGTGGTCCATGCTGCTAAACAAGACA	1634
Qy	1381	AAAGTCTGTCTTCCCTGAACCATGTGATAACTTTTACAGAAATGGACTCGAGCTCATCTG	1440
Db	1635	AAAGTCTGTCTTCCCTGAACCATGTGATAACTTTTACAGAAATGGACTCGAGCTCATCTG	1694
Qy	1441	CAAAAGGCCCTCTGTAAGACTGGTTTCTGCGCAATGACCAACAAGCCAAAGATTTTCCTC	1500
Db	1695	CAAAAGGCCCTCTGTAAGACTGGTTTCTGCGCAATGACCAACAAGCCAAAGATTTTCCTC	1754
Qy	1501	TTGTGATTTCTTTAAAAAGATGACTATATAATTTATTTCCACTAAAAATATGTTTCTGC	1560
Db	1755	TTGTGATTTCTTTAAAAAGATGACTATATAATTTATTTCCACTAAAAATATGTTTCTGC	1814
Qy	1561	ATTTCATTTTATAGCAACAATGTGTAACAACTCACTGTGATCAATATTTTATATCAT	1620
Db	1815	ATTTCATTTTATAGCAACAATGTGTAACAACTCACTGTGATCAATATTTTATATCAT	1874

Qy	1621	GCAAAATATGTTTAAATAAATAAGAAATTTGTAATTTATATAAAAAAAAAAAAAA	1674
Db	1875	GCAAAATATGTTTAAATAAATAAGAAATTTGTAATTTATATAAAAAAAAAAAAAA	1928
RESULT 6			
AR231261			
LOCUS	AR231261	2031 bp	linear
DEFINITION	Sequence 1 from patent US 6451764.		
ACCESSION	AR231261		
VERSION	AR231261.1	GI:27272182	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1. (bases 1 to 2031)		
TITLE	Lee, J. and Wood, W.		
JOURNAL	VEGF-related protein		
FEATURES	Patent: US 6451764-A 1 17-SBP-2002;		
source	Location/Qualifiers		
	1..2031		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
Query Match	99.4%;	Score 1664.6;	DB 6; Length 2031;
Best Local Similarity	99.8%;	Pred. No. 0;	
Matches 1667;	Conservative	0; Mismatches	4; Indels 0; Gaps
Qy	1	GTCTTCCACATGCACATCGCTGGCTTCTTCTCTGTGGGTGTTCTCTGCTCGCGCTG	60
Db	361	GTCTTCCACATGCACATCGCTGGCTTCTTCTCTGTGGGTGTTCTCTGCTCGCGCTG	420
Qy	61	CGTGTCTCCGGTCTCTCGAGCGCGCCGCCGCCGCCCTTCGAGTCCGGAATCG	120
Db	421	CGTGTCTCCGGTCTCTCGAGCGCGCCGCCGCCCTTCGAGTCCGGAATCG	480
Qy	121	ACCTCTCGACGCGGAGCCGACGCGGGGAGGCGACGCGCTTATGCAAGCAAGATCTCG	180
Db	481	ACCTCTCGACGCGGAGCCGACGCGGGGAGGCGACGCGCTTATGCAAGCAAGATCTCG	540
Qy	181	AGGAGCAGTTACCGTCTGTGTCCAGTGTAGATGAACATCATGCTGTACTCTACCCAGAAT	240
Db	541	AGGAGCAGTTACCGTCTGTGTCCAGTGTAGATGAACATCATGCTGTACTCTACCCAGAAT	600
Qy	241	ATTGGAATATGTAAGTGTCTAGCTAAGAAAGGAGCTGGCAACATAACAGAGAACAG	300
Db	601	ATTGGAATATGTAAGTGTCTAGCTAAGAAAGGAGCTGGCAACATAACAGAGAACAG	660
Qy	301	CCAACTCAACTCAAGGACAGAGACATATAAATTTTCTGCAGCACATTAATAACAG	360
Db	661	CCAACTCAACTCAAGGACAGAGACATATAAATTTTCTGCAGCACATTAATAACAG	720
Qy	361	AGATCTGAAAAGTATGTATAGTGTGAGAAAGACTCAATGATGCCACGGAGGTCT	420
Db	721	AGATCTGAAAAGTATGTATAGTGTGAGAAAGACTCAATGATGCCACGGAGGTCT	780
Qy	421	GTATAGATGTGGGAAGGAGTTTGAGTGTGCGACAAAACACCTTCTTTAACTCCATGTG	480
Db	781	GTATAGATGTGGGAAGGAGTTTGAGTGTGCGACAAAACACCTTCTTTAACTCCATGTG	840
Qy	481	TGTCCTGTACAGATGTGGGGTGTGTCATATGATGTAGGGGTGTGATGTCATGAACCA	540
Db	841	TGTCCTGTACAGATGTGGGGTGTGTCATATGATGTAGGGGTGTGATGTCATGAACCA	900
Qy	541	GCACGAGCTACCTCAGCAGACGTTATTGTAATTTACAGTGCTCTCTCTCAAGCCCCA	600
Db	901	GCACGAGCTACCTCAGCAGACGTTATTGTAATTTACAGTGCTCTCTCTCAAGCCCCA	960
Qy	601	AACAGTAAATATGTTTGGCAATCACTCTCTGCGGATGTCATGTCTAACTGGATG	660
Db	961	AACAGTAAATATGTTTGGCAATCACTCTCTGCGGATGTCATGTCTAACTGGATG	1020

QY 661 TTTACAGCAAGTTTCATTCCATTTATTAGACGTTCCCTGCCAGCAACTACCAAGTGTG 720
Db |||||
QY 1021 TTTACAGCAAGTTTCATTCCATTTATTAGACGTTCCCTGCCAGCAACTACCAAGTGTG 1080
Db |||||
QY 721 AGGCAGGCAACAGACCTGCCGCCCAATTAATACATGTGGAATAATCAATCTGCAGATGCC 780
Db |||||
QY 1081 AGGCAGGCAACAGACCTGCCGCCCAATTAATACATGTGGAATAATCAATCTGCAGATGCC 1140
Db |||||
QY 781 TGGCTCAGGAAGATTATTTATGTTTCTCGGATGCTGGAGATGATCAACAGATGGATTCC 840
Db |||||
QY 1141 TGGCTCAGGAAGATTATTTATGTTTCTCGGATGCTGGAGATGATCAACAGATGGATTCC 1200
Db |||||
QY 841 ATGACATCTGTGGACCAAAACAAGAGCTGGATGAAGAGACCTGTCACTGTGTCTGCAGAG 900
Db |||||
QY 1201 ATGACATCTGTGGACCAAAACAAGAGCTGGATGAAGAGACCTGTCACTGTGTCTGCAGAG 1260
Db |||||
QY 901 CGGGCTTCGGCTGCCAGCTGTGGACCCACAAAGAACTAGACAGAACTCATGCCAGT 960
Db |||||
QY 1261 CGGGCTTCGGCTGCCAGCTGTGGACCCACAAAGAACTAGACAGAACTCATGCCAGT 1320
Db |||||
QY 961 GTGTCTGTAAAAAACAACCTCTTCCCAAGCCAATGTGGGGCCAAACCGAGAATTTGATGAAA 1020
Db |||||
QY 1321 GTGTCTGTAAAAACAACCTCTTCCCAAGCCAATGTGGGGCCAAACCGAGAATTTGATGAAA 1380
Db |||||
QY 1021 ACATATGCCAGTGTGTATGTAAAGAACTGTGCCCAAGAAATCAACCCCTAAATCTCTGGAA 1080
Db |||||
QY 1381 ACATATGCCAGTGTGTATGTAAAGAACTGTGCCCAAGAAATCAACCCCTAAATCTCTGGAA 1440
Db |||||
QY 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAGAAAGTGTCC 1140
Db |||||
QY 1441 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAGAAAGTGTCC 1500
Db |||||
QY 1141 ACCAACAACATGCAGCTGTACAGAGCGGCATGTACGAACCGCCAGAAAGGTTGTGAGC 1200
Db |||||
QY 1501 ACCAACAACATGCAGCTGTACAGAGCGGCATGTACGAACCGCCAGAAAGGTTGTGAGC 1560
Db |||||
QY 1201 CAGGATTTTCATATAGTAAGAAAGTGTGCTGTTGCTCCCTCATATTTGGGAAAGACAC 1260
Db |||||
QY 1561 CAGGATTTTCATATAGTAAGAAAGTGTGCTGTTGCTCCCTCATATTTGGAAGAAAGACAC 1620
Db |||||
QY 1261 AAATGAGCTAAGATTGTACTGTTTTCAGTTTCATCGATTTTCTTATTTGGAAGAACTGTG 1320
Db |||||
QY 1621 AAATGAGCTAAGATTGTACTGTTTTCAGTTTCATCGATTTTCTTATTTGGAAGAACTGTG 1680
Db |||||
QY 1321 TGGCAACAGTAGAAGTGTCTGTGAACAGAGAGACCCCTGTGGGTCCATGTAAACAAAGACA 1380
Db |||||
QY 1681 TGGCAACAGTAGAAGTGTCTGTGAACAGAGAGACCCCTGTGGGTCCATGTAAACAAAGACA 1740
Db |||||
QY 1381 AAAGTCTGTCTTCCCTGAACCATGTGGATAAATTTACAGAAATGGACTGGAGCTCATCTG 1440
Db |||||
QY 1741 AAAGTCTGTCTTCCCTGAACCATGTGGATAAATTTACAGAAATGGACTGGAGCTCATCTG 1800
Db |||||
QY 1441 CAAAAGGCTCTTGTAAAGATGTTTCTGCCAATGACCAACAGCCCAAGATTTTCTCTC 1500
Db |||||
QY 1801 CAAAAGGCTCTTGTAAAGATGTTTCTGCCAATGACCAACAGCCCAAGATTTTCTCTC 1860
Db |||||
QY 1501 TTGTGATTTCTTTAAAGATGACTATATAATTTATTTCCACTTAAATATTTGTTTCTGCG 1560
Db |||||
QY 1861 TTGTGATTTCTTTAAAGATGACTATATAATTTATTTCCACTTAAATATTTGTTTCTGCG 1920
Db |||||
QY 1561 ATTCATTTTATAGCAACAACTTGGTAAACTCACTGTGATCAATATTTTATATATCAT 1620
Db |||||
QY 1921 ATTCATTTTATAGCAACAACTTGGTAAACTCACTGTGATCAATATTTTATATCAT 1980
Db |||||
QY 1621 GCAAAATATGTTTAAATTAATGAAATTTGATTTATTAATAAAAAA 1671
Db |||||
QY 1981 GCAAAATATGTTTAAATTAATGAAATTTGATTTATTAATAAAAAA 2031
Db |||||

RESULT 7
AR342826 2031 bp DNA linear PAT 17-AUG-2003
LOCUS

DEFINITION Sequence 1 from patent US 6576608.
ACCESSION AR342826
VERSION AR342826.1 GI:33738052
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 2031)
AUTHORS Lee,J. and Wood,W.
TITLE Methods of using VEGF-related protein
JOURNAL Patent: US 6576608-A 1 10-JUN-2003;
FEATURES Location/Qualifiers
1..2031
source /organism="unknown"
ORIGIN /mol_type="genomic DNA"

Query Match 99.4%; Score 1664.6; DB 6; Length 2031;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1667; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GTCCCTTCCACCATGCACCTCGCTGGGCTTCTTCTCTGTGGGGTGTCTCTGTCTGCCCGCTG 60
Db GTCCCTTCCACCATGCACCTCGCTGGGCTTCTTCTCTGTGGGGTGTCTCTGTCTGCCCGCTG 420
QY 61 CGCTGCTCCCGGTCCTCGCGAGCGCCGCGCGCGCGCGCTTCGAGTCCCGGACTCG 120
Db CGCTGCTCCCGGTCCTCGCGAGCGCGCGCGCGCGCGCTTCGAGTCCCGGACTCG 480
QY 121 ACCTCTGGAGCGCGAGCGCCGACCGCGGCGAGGCCACGGCTTATGGAAGAAAGATCTGG 180
Db ACCTCTGGAGCGCGAGCGCCGACCGCGGCGAGGCCACGGCTTATGGAAGAAAGATCTGG 540
QY 181 AGGAGCAGTGTACGCTGTGTGCCAGTGTAGATGAATCATGACTGATCTCTTACCCAGAAT 240
Db AGGAGCAGTGTACGCTGTGTGCCAGTGTAGATGAATCATGACTGATCTCTTACCCAGAAT 600
QY 241 ATTGGAAAATGTACAAAGTGCAGCTAAAGAAAGAGGCTGCGCAACATAACAGAGAAACAGG 300
Db ATTGGAAAATGTACAAAGTGCAGCTAAAGAAAGAGGCTGCGCAACATAACAGAGAAACAGG 660
QY 301 CCAACCTCACTCAAGACACAGAGAGACTATAAATTTGCTGCGAGACATTATAATACAG 360
Db CCAACCTCACTCAAGACACAGAGAGACTATAAATTTGCTGCGAGACATTATAATACAG 720
QY 361 AGATCTTTGAAAAGTATTGTATAATGAGTGGAGAAAGACTCAATGCATGCCACGGAGGTGT 420
Db AGATCTTTGAAAAGTATTGTATAATGAGTGGAGAAAGACTCAATGCATGCCACGGAGGTGT 780
QY 421 GTATAGATGTGGGAGAGAGTGTGGAGTCGGGCAAAACACCTCTTTTAAACCTCCATGTG 480
Db GTATAGATGTGGGAGAGAGTGTGGAGTCGGGCAAAACACCTCTTTTAAACCTCCATGTG 840
QY 481 TGTCCGTCTACAGATGTGGGGGTGTGCTCAATAGTGGGGGCTGCAGTGCATGAACACCA 540
Db TGTCCGTCTACAGATGTGGGGGTGTGCTCAATAGTGGGGGCTGCAGTGCATGAACACCA 900
QY 541 GCACAGCTACCTCAGCAAGACGTTATTGAAATTTACAGTGCCTCTCTCTCAAGGGCCCCA 600
Db GCACAGCTACCTCAGCAAGACGTTATTGAAATTTACAGTGCCTCTCTCTCAAGGGCCCCA 960
QY 601 AACCAGTAAACAATCAGTTTTTGCACATCACAATCTTCCTGCCGATGCATGTCTTAACTGGATG 660
Db AACCAGTAAACAATCAGTTTTTGCACATCACAATCTTCCTGCCGATGCATGTCTTAACTGGATG 1020
QY 661 TTTACAGCAAGTTTCATTCCATTTATTAGAGCTTCCCTGCCAGCAACACTACCAAGTGTG 720
Db TTTACAGCAAGTTTCATTCCATTTATTAGAGCTTCCCTGCCAGCAACACTACCAAGTGTG 1080
QY 721 AGGCAGGCAACAGACCTGCCCCACCAATTTACATGTGGGAATAATCATCTGCAGATGCC 780
Db AGGCAGGCAACAGACCTGCCCCACCAATTTACATGTGGGAATAATCATCTGCAGATGCC 1140

```

QY 781 TGGCTCAGAGATTTTATGTTTCTCGATGCTGGAGATGATCTCAACAGATGATCC 840
Db 1141 TGGCTCAGAGATTTTATGTTTCTCGATGCTGGAGATGATCTCAACAGATGATCC 1200
QY 841 ATGACATCTGTGGACCAACCAAGAGCTGGATGAAGAGACCTGTCAGTGTCTGCGAGAG 900
Db 1201 ATGACATCTGTGGACCAACCAAGAGCTGGATGAAGAGACCTGTCAGTGTCTGCGAGAG 1260
QY 901 CGGGCTTTGGCTGCCAGCTGTGGACCCCAAGAACCTAGACAGAACTCATGCCAGT 960
Db 1261 CGGGCTTTGGCTGCCAGCTGTGGACCCCAAGAACCTAGACAGAACTCATGCCAGT 1320
QY 961 GTGCTGTAAACCAACCTCTTCCCGCCCAATGCTGGGCCCAACCGGAATTTGATGAAA 1020
Db 1321 GTGCTGTAAACCAACCTCTTCCCGCCCAATGCTGGGCCCAACCGGAATTTGATGAAA 1380
QY 1021 ACACATGCCAGCTGTATGTAAGAGACCTGCCGCCAGAAATCAACCCCTAAATCCTCGAA 1080
Db 1381 ACACATGCCAGCTGTATGTAAGAGACCTGCCGCCAGAAATCAACCCCTAAATCCTCGAA 1440
QY 1081 AATGTCCTGTGAATGTACAGAAAGTCCAAGAAATGCTTGTAAAGAGAAAGATTC 1140
Db 1441 AATGTCCTGTGAATGTACAGAAAGTCCAAGAAATGCTTGTAAAGAGAAAGATTC 1500
QY 1141 ACCACCAACATGAGCTGTACAGACGGCCATGTACGACCGCCAGAGCTTTGTGAGC 1200
Db 1501 ACCACCAACATGAGCTGTACAGACGGCCATGTACGACCGCCAGAGCTTTGTGAGC 1560
QY 1201 CAGGATTTTCATATAGTGAAGAGTGTGCTGCTGCTCCCTCATATTTGGCAAGACCAC 1260
Db 1561 CAGGATTTTCATATAGTGAAGAGTGTGCTGCTGCTCCCTCATATTTGGCAAGACCAC 1620
QY 1261 AAATGAGCTAAGATTGACTGTTTCCAGTTCATCGATTTCTATTATGAAAGCTGTGT 1320
Db 1621 AAATGAGCTAAGATTGACTGTTTCCAGTTCATCGATTTCTATTATGAAAGCTGTGT 1680
QY 1321 TGCCACAGTAGAAGCTGTGTGAACAGAGAGACCTTTGGGTCCATGCTTAACAAAGACA 1380
Db 1681 TGCCACAGTAGAAGCTGTGTGAACAGAGAGACCTTTGGGTCCATGCTTAACAAAGACA 1740
QY 1381 AAATGCTGCTTCTTCCAGCATGTAATACATTTTACAGAAATGAGTGGAGCTCATCTG 1440
Db 1741 AAATGCTGCTTCTTCCAGCATGTAATACATTTTACAGAAATGAGTGGAGCTCATCTG 1800
QY 1441 CAAAGGCTCTTGTAAAGACTGTTTCTGCCAATGACCAACAGCCAGAGATTTTCCCTC 1500
Db 1801 CAAAGGCTCTTGTAAAGACTGTTTCTGCCAATGACCAACAGCCAGAGATTTTCCCTC 1860
QY 1501 TTGTGATTTCTTTAAAGAGATGACTATATAATTTTATTTTCCACTAAATAATTTCTTCTGC 1560
Db 1861 TTGTGATTTCTTTAAAGAGATGACTATATAATTTTATTTTCCACTAAATAATTTCTTCTGC 1920
QY 1561 ATTCAATTTTATAGCAACAAATTTGTAAGACTGCTGTAAGTCAATATTTTATATCAT 1620
Db 1921 ATTCAATTTTATAGCAACAAATTTGTAAGACTGCTGTAAGTCAATATTTTATATCAT 1980
QY 1621 GCAAAATATGTTTAAATAAATAAATGAAATTTGATTTTATAAAAAA 1671
Db 1981 GCAAAATATGTTTAAATAAATAAATGAAATTTGATTTTATAAAAAA 2031

RESULT 8
BC063685 LOCUS
DEFINITION Homo sapiens cDNA clone MGC:74703 IMAGE:4513934, complete cds.
ACCESSION BC063685
VERSION BC063685.1 GI:39645812
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2076)

```

AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altshuler, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.P., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, E.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Scaplehorn, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Carninci, P., Brownstein, M.J., Udwin, T.B., Toshiyuki, S.,
 Schmitt, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Rahy, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S.,
 Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywicki, M.I., Skalska, U., Smalls, D.E.,
 Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 2076)
 Strausberg, R.
 Direct Submission
 Submitted (08-DEC-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@nigri.nih.gov
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Lalic, P., Legaspi, R.,
 Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 134 Row: j Column: 2
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 19924300.
 Location/Qualifiers
 1..2076
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:74703 IMAGE:4513934"
 /tissue_type="Bladder, transitional cell papilloma"
 /clone_lib="NIH MGC_93"
 /lab_host="DH10B"
 /note="vector: pCMV-SPORT6"
 417..1676
 /codon_start=1
 /product="unknown (protein for MGC:74703)"
 /protein_id="AAH63685.1"
 /db_xref="GI:39645813"
 /translation="MHLLGFFSVACSLAAALLPGPREFAFAAFAAAGSLDSDAEPD
 AGEATAYASKDLERQLRSVSDLMVTLYPEYWKMKCOLRKGQWQHREQALNSR
 TEETIKGAAAHYNTILKSIDNEWKTCQMPREVCIDVGEFVATNTFFKPCVSVY
 RCGGCCNSEGQCNCNTSLSKTLFELTPTLSQGPVKPTISFANHTSCRCMSKLDVY

FEATURES

source

CDS

RVHSIIIRSLPATLPOCOAANKTCPTNYMNNHICRCLAQEDFESSDAGDSDTGGF
HDICGNKLEDETCOCVCPAGIIPASCGPHKELDENSCOCUCNKLIPSOQGANRFP
DENTCOVCKETCPNPNPGKACACECTESPOKLLKXKHQHTCSCIRPFCNRR
KACEPFSYSEEVCRVPSWYKSPQMS"
801..1055
/note="PDGF; Region: Platelet-derived and vascular
endothelial growth factors (PDGF, VEGF) family"
/db_xref="CDD:smart00141"

misc_feature

ORIGIN

Query Match 99.4%; Score 1664.6; DB 9; Length 2076;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1667; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GTCTCTCCACCATGCACTCGCTGGGCTTCTCTCTGTGGCGTGTCTCTCGCGCTG 60
DB 406 GTCTCTCCACCATGCACTGCTGGGCTTCTCTCTGTGGCGTGTCTCTCGCGCTG 465
QY 61 CGTGTCTCCCGGTCTCTCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 120
DB 466 CGTGTCTCCCGGTCTCTCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 525
QY 121 ACCTCTCGGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 180
DB 526 ACCTCTCGGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 585
QY 181 AGGAGCAGTTACCGTCTGTGTCCAGTGTAGTGAATCATGACTGTACTTACCCAGAT 240
DB 586 AGGAGCAGTTACCGTCTGTGTCCAGTGTAGTGAATCATGACTGTACTTACCCAGAT 645
QY 241 ATTGGAATATGACAGTGTGCTAAGGAAGAGGCTGGCAACATAACAGAGAACAGG 300
DB 646 ATTGGAATATGACAGTGTGCTAAGGAAGAGGCTGGCAACATAACAGAGAACAGG 705
QY 301 CCAACCTCAACTCAAGGACAGAGAGACTATAAATTTGCTGCAGCACATTAATACAG 360
DB 706 CCAACCTCAACTCAAGGACAGAGAGACTATAAATTTGCTGCAGCACATTAATACAG 765
QY 361 AGATCTTGAAGATATTGATAGTGGAGAAAGACTCAATGCAATGCCAGGGAGGTGT 420
DB 766 AGATCTTGAAGATATTGATAGTGGAGAAAGACTCAATGCAATGCCAGGGAGGTGT 825
QY 421 GTATAGATGTGGGAAGGAGTTTGGAGTCGCGACAAACACCTTCTTTAACTCCATGTG 480
DB 826 GTATAGATGTGGGAAGGAGTTTGGAGTCGCGACAAACACCTTCTTTAACTCCATGTG 885
QY 481 TGTCCGTCTACAGATGTGGGGTGTGTCGAATAGTGGGGGTGAGTGCATGAACACCA 540
DB 886 TGTCCGTCTACAGATGTGGGGTGTGTCGAATAGTGGGGGTGAGTGCATGAACACCA 945
QY 541 GCACGAGCTACCTCAGCAGAGAGTTATTGAAATTTACAGTGCCTCTCTCAAGGCCCA 600
DB 946 GCACGAGCTACCTCAGCAGAGAGTTATTGAAATTTACAGTGCCTCTCTCAAGGCCCA 1005
QY 601 AACCACTAAACATCAGTTTGGCAATCACACTTCTCGCGATGCATGTCTAACTGGATG 660
DB 1006 AACCACTAAACATCAGTTTGGCAATCACACTTCTCGCGATGCATGTCTAACTGGATG 1065
QY 661 TTACAGACAGTTTATTCATATTAGACGTTCCTCGCGAGCAACACTACACAGTGTG 720
DB 1066 TTACAGACAGTTTATTCATATTAGACGTTCCTCGCGAGCAACACTACACAGTGTG 1125
QY 721 AGGACGGAACAGACTCCGCCACCAATTACATGTGGAATAATCACATCTGCAGATGC 780
DB 1126 AGGACGGAACAGACTCCGCCACCAATTACATGTGGAATAATCACATCTGCAGATGC 1185
QY 781 TGGCTCAGGAAGATTTATGTTTCTCGGATGCTGGAGTGAATCAACAGATGGAATCC 840
DB 1186 TGGCTCAGGAAGATTTATGTTTCTCGGATGCTGGAGTGAATCAACAGATGGAATCC 1245
QY 841 ATGACATCTGTGGACCAACAGGAGCTGGATGAAGAGACCTGTGCTGTCTGCAGAG 900
DB 1246 ATGACATCTGTGGACCAACAGGAGCTGGATGAAGAGACCTGTGCTGTCTGCAGAG 1305

QY 901 CGGGGCTTGGCCCTGCCAGCTGTGGACCCCAAGAACTAGACAGAACTCATGCCAGT 960
DB 1306 CGGGGCTTGGCCCTGCCAGCTGTGGACCCCAAGAACTAGACAGAACTCATGCCAGT 1365
QY 961 GTGCTGTAAACAACTCTTCCCGAGCCAACTGTGGGCGCAACCGAGAAATTTTCATGAA 1020
DB 1366 GTGCTGTAAACAACTCTTCCCGAGCCAACTGTGGGCGCAACCGAGAAATTTTCATGAA 1425
QY 1021 ACACATGCCAGTGTGTATGTAAAGAACTGCGCCAGAAATCAACCCCTAAATCCTGGAA 1080
DB 1426 ACACATGCCAGTGTGTATGTAAAGAACTGCGCCAGAAATCAACCCCTAAATCCTGGAA 1485
QY 1081 AATGCTGTGTATGTACAGAAAGTCCACAGAAATGCTTGTAAAGAAAGAGAGTTC 1140
DB 1486 AATGCTGTGTATGTACAGAAAGTCCACAGAAATGCTTGTAAAGAAAGAGAGTTC 1545
QY 1141 ACCACCAACATGAGCTGTGTACAGAGCGGCATGTACGAACCGCCAGAAAGCTTGTGAGC 1200
DB 1546 ACCACCAACATGAGCTGTGTACAGAGCGGCATGTACGAACCGCCAGAAAGCTTGTGAGC 1605
QY 1201 CAGGATTTTCATATAGTGAAGAGTGTGCTGTGTGCTCCCTTCATATTTGGCAAGACCAC 1260
DB 1606 CAGGATTTTCATATAGTGAAGAGTGTGCTGTGTGCTCCCTTCATATTTGAAAAGACCAC 1665
QY 1261 AATGAGCTAAGATTGACTGTGTTTCCAGTTTCATCGATTTTCTATTATGAAAAGCTGTG 1320
DB 1666 AATGAGCTAAGATTGACTGTGTTTCCAGTTTCATCGATTTTCTATTATGAAAAGCTGTG 1725
QY 1321 TGCCACAGTGAAGTGTGTGTGAACAGAGAGACCCCTTGTGGTCCATGCTTAAACAAAGACA 1380
DB 1726 TGCCACAGTGAAGTGTGTGTGAACAGAGAGACCCCTTGTGGTCCATGCTTAAACAAAGACA 1785
QY 1381 AAGTCTGTCTTCTGCAACCATGTGTGATAACTTTACAGAAATGGAGTGGAGTCTCATCTG 1440
DB 1786 AAGTCTGTCTTCTGCAACCATGTGTGATAACTTTACAGAAATGGAGTGGAGTCTCATCTG 1845
QY 1441 CAAAGGCCCTTGTGAAGACTGGTTTTCTGCCAATGACCAACAGCAAGAGATTTTCTC 1500
DB 1846 CAAAGGCCCTTGTGAAGACTGGTTTTCTGCCAATGACCAACAGCAAGAGATTTTCTC 1905
QY 1501 TTGTGATTTCTTTAAAGAGTGAATTAATTTTATTTTCCACTAAATAATTTTCTTCTG 1560
DB 1906 TTGTGATTTCTTTAAAGAGTGAATTAATTTTATTTTCCACTAAATAATTTTCTTCTG 1965
QY 1561 ATTCAATTTTATAGCAACAACTTGTGTAATCTCACTGTGATCAATATTTTATATCAT 1620
DB 1966 ATTCAATTTTATAGCAACAACTTGTGTAATCTCACTGTGATCAATATTTTATATCAT 2025
QY 1621 GCAAAATATGTTTAAATAAATGAAATGTAATTTTATATAAAAAA 1671
DB 2026 GCAAAATATGTTTAAATAAATGAAATGTAATTTTATATAAAAAA 2076

RESULT 9

AR231262/c

LOCUS

DEFINITION

AR231262

SEQUENCE

AR231262.1

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

SOURCE

ORIGIN

Sequence 2 from patent US 6451764.
GI:27272183

Unknown.
Unclassified.
1 (bases 1 to 2031)
Lee, J. and Wood, W.
VEGF-related protein
Patent: US 6451764-A 2 17-SEP-2002;
Location/Qualifiers
1..2031
/organism="unknown"
/mol_type="genomic DNA"

AR231262 2031 bp DNA linear PAT 20-DEC-2002

Qy	1	GTCTTCCACCATGACCTGGCTTCTCTCTGTCGCGTGTCTCTGTCGCCGCTG	60
Db	1671	GTCTTCCACCATGACCTGGCTTCTCTCTGTCGCGTGTCTCTGTCGCCGCTG	1612
Qy	61	CGTGTCTCCGGTCTCTGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	120
Db	1611	TGCTGTCTCCGGTCTCTGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1552
Qy	121	ACCTCTCGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	180
Db	1551	ACCTCTCGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1492
Qy	181	AGGAGCAGTTACGCTGTGTGTCAGTGTAGATGAATCATGACTGTACTCTACCCAGAA	240
Db	1491	AGGAGCAGTTACGCTGTGTGTCAGTGTAGATGAATCATGACTGTACTCTACCCAGAA	1432
Qy	241	ATTGGAAATCTACAAGTGTCTGCTAGGTAAGGAGGCTGGCAACATAACAGAGACAGG	300
Db	1431	ATTGGAAATCTACAAGTGTCTGCTAGGTAAGGAGGCTGGCAACATAACAGAGACAGG	1372
Qy	301	CCAACTCTCAAGCAGAGAGAGACTATAAAATTTGCTGAGCAGCAATTAATAACAG	360
Db	1371	CCAACTCTCAAGCAGAGAGAGACTATAAAATTTGCTGAGCAGCAATTAATAACAG	1312
Qy	361	AGATCTTGAAAGTATTGATAATGAGTGGAGAAAGACTCAATGATGCCACGGAGGTG	420
Db	1311	AGATCTTGAAAGTATTGATAATGAGTGGAGAAAGACTCAATGATGCCACGGAGGTG	1252
Qy	421	GTATAGATGTGGGAAAGAGTTTGGAGTTCGCGACCAACACTCTCTTTAAACCTCCATGTG	480
Db	1251	GTATAGATGTGGGAAAGAGTTTGGAGTTCGCGACCAACACTCTCTTTAAACCTCCATGTG	1192
Qy	481	TGTCGCTCTACAGATGTGGGGTTTGTGCAATAGTAGGGGCTGCAATGATGAACACCA	540
Db	1191	TGTCGCTCTACAGATGTGGGGTTTGTGCAATAGTAGGGGCTGCAATGATGAACACCA	1132
Qy	541	GCAGAGCTACCTCAGCAAGAGCTTATTGAATTAAGTGCCTCTCTCAGAGGCCCA	600
Db	1131	GCAGAGCTACCTCAGCAAGAGCTTATTGAATTAAGTGCCTCTCTCAGAGGCCCA	1072
Qy	601	AACAGTAAACATCAGTCTTTCGCAATCACAATCTCTCCGATGCATGTCTAACTGGATG	660
Db	1071	AACAGTAAACATCAGTCTTTCGCAATCACAATCTCTCCGATGCATGTCTAACTGGATG	1012
Qy	661	TTTACAGACAGTTCATTCATTTATAGAGTTCCTGCGAGTCAACACAGATGATCC	720
Db	1011	TTTACAGACAGTTCATTCATTTATAGAGTTCCTGCGAGTCAACACAGATGATCC	952
Qy	721	AGGAGCGAACAAGACCTGCCACCAATTAATGATGGAATAATCATCTGCAGATGCC	780
Db	951	AGGAGCGAACAAGACCTGCCACCAATTAATGATGGAATAATCATCTGCAGATGCC	892
Qy	781	TGGCTCAGGAAGATTTTATGTTTCTCGAGTCTGAGATGACTCAACAGATGATCC	840
Db	891	TGGCTCAGGAAGATTTTATGTTTCTCGAGTCTGAGATGACTCAACAGATGATCC	832
Qy	841	ATGACATCTGCGACCAACAAGAGCTGTGATGAAGAGCTGTGATGTCTGCAGAG	900
Db	831	ATGACATCTGCGACCAACAAGAGCTGTGATGAAGAGCTGTGATGTCTGCAGAG	772
Qy	901	CGGGCTTCGGCTGCGAGCTGGAGCCCAACAAAGACTAGACAAACTCATGCGAGT	960
Db	771	CGGGCTTCGGCTGCGAGCTGGAGCCCAACAAAGACTAGACAAACTCATGCGAGT	712
Qy	961	GTGTCTGTAACAAACTCTTCCGACGCAATGTGTGGGCGCAACCGAGAAATTTGATGAA	1020
Db	711	GTGTCTGTAACAAACTCTTCCGACGCAATGTGTGGGCGCAACCGAGAAATTTGATGAA	652
Qy	1021	ACACATGCCAGTGTGTATGTAAAGAACCTGCCCCAGAAATCAACCCCTAAATCTTGAA	1080
Db	651	ACACATGCCAGTGTGTATGTAAAGAACCTGCCCCAGAAATCAACCCCTAAATCTTGAA	592
Qy	1081	AATGTCCCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAGAGTTCC	1140
Db	591	AATGTCCCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAGAGTTCC	532
Qy	1141	ACCACCAACATGACGCTGTTACAGACGCGCATGTACGAACCGCCAGAGGCTTGTGAGC	1200
Db	531	ACCACCAACATGACGCTGTTACAGACGCGCATGTACGAACCGCCAGAGGCTTGTGAGC	472
Qy	1201	CAGGATTTTCATATAGTGAAGAGTGTGCTGTTGTGCCCTTCATATGGCAAGACAC	1260
Db	471	CAGGATTTTCATATAGTGAAGAGTGTGCTGTTGTGCCCTTCATATGGCAAGACAC	412
Qy	1261	AAATGAGCTTAAGATTGTACTGTTTCCAGTTTCATCGATTTTCTATATGGAATACTGTG	1320
Db	411	AAATGAGCTTAAGATTGTACTGTTTCCAGTTTCATCGATTTTCTATATGGAATACTGTG	352
Qy	1321	TGCCACAGTAGAATCTGTCTGTGAACAGAGAGACCTTGTGGTCCATGCTAACAAAGACA	1380
Db	351	TGCCACAGTAGAATCTGTCTGTGAACAGAGAGACCTTGTGGTCCATGCTAACAAAGACA	292
Qy	1381	AAAGTCTGTCTTCTCTGACCACTGTGGATACTTTTACAGAAATGACTGGAGCTCATCTG	1440
Db	291	AAAGTCTGTCTTCTCTGACCACTGTGGATACTTTTACAGAAATGACTGGAGCTCATCTG	232
Qy	1441	CAAAAGGCTCTTGTAAAGACTGGTTTCTGCCAATGACCAACCAAGATTTTCTCTC	1500
Db	231	CAAAAGGCTCTTGTAAAGACTGGTTTCTGCCAATGACCAACCAAGATTTTCTCTC	172
Qy	1501	TTGTGATTTCTTAAAGAAATGACTATATAATTTTCCACTAAATAATTTGTTCTGC	1560
Db	171	TTGTGATTTCTTAAAGAAATGACTATATAATTTTCCACTAAATAATTTGTTCTGC	112
Qy	1561	ATTCAATTTTATAGCAACAACATTTGGTAAACTCCTGTGATCAATATTTTATATCAT	1620
Db	111	ATTCAATTTTATAGCAACAACATTTGGTAAACTCCTGTGATCAATATTTTATATCAT	52
Qy	1621	GCAAAATATGTTTAAATAAATAAATGAAATGATTTATATAAAAAA	1671
Db	51	GCAAAATATGTTTAAATAAATAAATGAAATGATTTATATAAAAAA	1
<p>RESULT 10</p> <p>AR342827/c</p> <p>LOCUS AR342827</p> <p>DEFINITION Sequence 2 from patent US 6576608.</p> <p>ACCESSION AR342827</p> <p>VERSION AR342827.1</p> <p>KEYWORDS GI:33738053</p> <p>SOURCE Unknown.</p> <p>ORGANISM Unclassified.</p> <p>REFERENCE 1 (bases 1 to 2031)</p> <p>AUTHORS Lee, J. and Wood, W.</p> <p>TITLE Methods of using VEGF-related protein</p> <p>JOURNAL Patent: US 6576608-A 2 10-JUN-2003;</p> <p>FEATURES</p> <p>Location/Qualifiers</p> <p>1..2031</p> <p>/organism="unknown"</p> <p>/mol_type="genomic DNA"</p>			
Qy	1	GTCTTCCACCATGACCTGGCTTCTCTCTGTCGCGTGTCTCTGTCGCCGCTG	60
Db	1671	GTCTTCCACCATGACCTGGCTTCTCTCTGTCGCGTGTCTCTGTCGCCGCTG	1612
Qy	61	CGTGTCTCCGGTCTCTGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	120
Db	1611	TGCTGTCTCCGGTCTCTGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1552
Qy	121	ACCTCTCGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	180
Db	1551	ACCTCTCGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1492
Qy	181	AGGAGCAGTTACGCTGTGTGTCAGTGTAGATGAATCATGACTGTACTCTACCCAGAA	240
Db	1491	AGGAGCAGTTACGCTGTGTGTCAGTGTAGATGAATCATGACTGTACTCTACCCAGAA	1432
Qy	241	ATTGGAAATCTACAAGTGTCTGCTAGGTAAGGAGGCTGGCAACATAACAGAGACAGG	300
Db	1431	ATTGGAAATCTACAAGTGTCTGCTAGGTAAGGAGGCTGGCAACATAACAGAGACAGG	1372
Qy	301	CCAACTCTCAAGCAGAGAGAGACTATAAAATTTGCTGAGCAGCAATTAATAACAG	360
Db	1371	CCAACTCTCAAGCAGAGAGAGACTATAAAATTTGCTGAGCAGCAATTAATAACAG	1312
Qy	361	AGATCTTGAAAGTATTGATAATGAGTGGAGAAAGACTCAATGATGCCACGGAGGTG	420
Db	1311	AGATCTTGAAAGTATTGATAATGAGTGGAGAAAGACTCAATGATGCCACGGAGGTG	1252
Qy	421	GTATAGATGTGGGAAAGAGTTTGGAGTTCGCGACCAACACTCTCTTTAAACCTCCATGTG	480
Db	1251	GTATAGATGTGGGAAAGAGTTTGGAGTTCGCGACCAACACTCTCTTTAAACCTCCATGTG	1192
Qy	481	TGTCGCTCTACAGATGTGGGGTTTGTGCAATAGTAGGGGCTGCAATGATGAACACCA	540
Db	1191	TGTCGCTCTACAGATGTGGGGTTTGTGCAATAGTAGGGGCTGCAATGATGAACACCA	1132
Qy	541	GCAGAGCTACCTCAGCAAGAGCTTATTGAATTAAGTGCCTCTCTCAGAGGCCCA	600
Db	1131	GCAGAGCTACCTCAGCAAGAGCTTATTGAATTAAGTGCCTCTCTCAGAGGCCCA	1072
Qy	601	AACAGTAAACATCAGTCTTTCGCAATCACAATCTCTCCGATGCATGTCTAACTGGATG	660
Db	1071	AACAGTAAACATCAGTCTTTCGCAATCACAATCTCTCCGATGCATGTCTAACTGGATG	1012
Qy	661	TTTACAGACAGTTCATTCATTTATAGAGTTCCTGCGAGTCAACACAGATGATCC	720
Db	1011	TTTACAGACAGTTCATTCATTTATAGAGTTCCTGCGAGTCAACACAGATGATCC	952
Qy	721	AGGAGCGAACAAGACCTGCCACCAATTAATGATGGAATAATCATCTGCAGATGCC	780
Db	951	AGGAGCGAACAAGACCTGCCACCAATTAATGATGGAATAATCATCTGCAGATGCC	892
Qy	781	TGGCTCAGGAAGATTTTATGTTTCTCGAGTCTGAGATGACTCAACAGATGATCC	840
Db	891	TGGCTCAGGAAGATTTTATGTTTCTCGAGTCTGAGATGACTCAACAGATGATCC	832
Qy	841	ATGACATCTGCGACCAACAAGAGCTGTGATGAAGAGCTGTGATGTCTGCAGAG	900
Db	831	ATGACATCTGCGACCAACAAGAGCTGTGATGAAGAGCTGTGATGTCTGCAGAG	772
Qy	901	CGGGCTTCGGCTGCGAGCTGGAGCCCAACAAAGACTAGACAAACTCATGCGAGT	960
Db	771	CGGGCTTCGGCTGCGAGCTGGAGCCCAACAAAGACTAGACAAACTCATGCGAGT	712
Qy	961	GTGTCTGTAACAAACTCTTCCGACGCAATGTGTGGGCGCAACCGAGAAATTTGATGAA	1020
Db	711	GTGTCTGTAACAAACTCTTCCGACGCAATGTGTGGGCGCAACCGAGAAATTTGATGAA	652

```
QY 61 CGTGTCTCCGGGTCTCTCGAGGCGCCCGCGCGCGCTTCGAGTCGGACTCG 120
Db 1611 TGTGTCTCCGGGTCTCTCGAGGCGCCCGCGCGCGCTTCGAGTCGGACTCG 1552
QY 121 ACCTCTCGGACGGGAGCCCGACGCGGGGAGGCCACGGCTTATGCAAGCAAGATCTCG 180
Db 1551 ACCTCTCGGACGGGAGCCCGACGCGGGGAGGCCACGGCTTATGCAAGCAAGATCTCG 1492
QY 181 AGAGCAGTTACGGTCTGTGTCCAGTGTAGATGAATCATGACTGTACTCTACCCAGAA 240
Db 1491 AGAGCAGTTACGGTCTGTGTCCAGTGTAGATGAATCATGACTGTACTCTACCCAGAA 1432
QY 241 ATTGGAATATGTAACAGTGTACGTAAGGAAGAGCGCTGGCAACATAACAGAGAACAG 300
Db 1431 ATTGGAATATGTAACAGTGTACGTAAGGAAGAGCGCTGGCAACATAACAGAGAACAG 1372
QY 301 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGTGTCAGACATTAATATACAG 360
Db 1371 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGTGTCAGACATTAATATACAG 1312
QY 361 AGATCTTGAAGTATTGATGATGAGTGGAGAGAGCTCAATGCTGCCACGGAGGTGT 420
Db 1311 AGATCTTGAAGTATTGATGATGAGTGGAGAGAGCTCAATGCTGCCACGGAGGTGT 1252
QY 421 GTATAGATGTGGGAGAGAGCTTTGAGTCCGCAACAAACACCTTCTTTAAACCTCCATGTG 480
Db 1251 GTATAGATGTGGGAGAGAGCTTTGAGTCCGCAACAAACACCTTCTTTAAACCTCCATGTG 1192
QY 481 TGTCCGTCTACAGATGTGGGGTGTCTGCAATAGTGAAGGCTGCAATGCAATGCAATGCA 540
Db 1191 TGTCCGTCTACAGATGTGGGGTGTCTGCAATAGTGAAGGCTGCAATGCAATGCAATGCA 1132
QY 541 GCAGAGCTACCTCAGCAAGACGTTATTGAAATTTACAGTGCCTCTCTCAAGGCCCA 600
Db 1131 GCAGAGCTACCTCAGCAAGACGTTATTGAAATTTACAGTGCCTCTCTCAAGGCCCA 1072
QY 601 AACAGTAAACATGAGTTTGCACATCACACTTCTTCCGATGATGCTTAACTGGATG 660
Db 1071 AACAGTAAACATGAGTTTGCACATCACACTTCTTCCGATGATGCTTAACTGGATG 1012
QY 661 TTTACAGACAGTTCATTCATTTAGAGTTCCTGCGAGGCTGCGAGCAACACTACACAGTGC 720
Db 1011 TTTACAGACAGTTCATTCATTTAGAGTTCCTGCGAGGCTGCGAGCAACACTACACAGTGC 952
QY 721 AGCAGCGAACAGAGCTGCCCAACCAATACATGTGGATATATCATCTCAGATGCC 780
Db 951 AGCAGCGAACAGAGCTGCCCAACCAATACATGTGGATATATCATCTCAGATGCC 892
QY 781 TGGCTCAGGAGATTTTATGTTTCTCGATGCTGGAGTGAATGACTCAACAGATGGATTC 840
Db 891 TGGCTCAGGAGATTTTATGTTTCTCGATGCTGGAGTGAATGACTCAACAGATGGATTC 832
QY 841 ATGACATCTGTGGACCAACAGAGAGCTGGATGAAGAGACCTCTCAGTGTGTCTGCAGAG 900
Db 831 ATGACATCTGTGGACCAACAGAGAGCTGGATGAAGAGACCTCTCAGTGTGTCTGCAGAG 772
QY 901 CGGGCTTCGGCTCGCAGCTGTGGACCCCAACAAAGAACTAGACAGAAATCTATGCCAGT 960
Db 771 CGGGCTTCGGCTCGCAGCTGTGGACCCCAACAAAGAACTAGACAGAAATCTATGCCAGT 712
QY 961 GTGTCTGTAACCAAACTCTTCCCAAGCAATGTGGGGCCCAACGAGAAATTTGATGAAA 1020
Db 711 GTGTCTGTAACCAAACTCTTCCCAAGCAATGTGGGGCCCAACGAGAAATTTGATGAAA 552
QY 1021 ACATGCCAGTGTGTATGTAAGAACTCTGCCGCCAGAAATCAACCCCTAAATCTCGGAA 1080
Db 651 ACATGCCAGTGTGTATGTAAGAACTCTGCCGCCAGAAATCAACCCCTAAATCTCGGAA 592
QY 1081 AATGTCTGTGAATGTACAGAGTCCACAGAAATGCTTGTGTAAGGAAGAGAGTCTCC 1140
Db 591 AATGTCTGTGAATGTACAGAGTCCACAGAAATGCTTGTGTAAGGAAGAGAGTCTCC 532
QY 1141 ACCACCAACATGAGCTGTATACAGAGCGGCTGTATGAGCGGCTTGTGAGC 1200
```

```
Db 531 ACCACCAACATGAGCTGTATACAGCGGCATGTACGAACCGCAGAGCTTGTGAGC 472
QY 1201 CAGGATTTTTCATATAGTGAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
Db 471 CAGGATTTTTCATATAGTGAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 412
QY 1261 AATGAGCTAAGATGTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320
Db 411 AATGAGCTAAGATGTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 352
QY 1321 TGCCACAGTAGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1380
Db 351 TGCCACAGTAGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 292
QY 1381 AAAGTGTGTCTTCTCTGAAACCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440
Db 291 AAAGTGTGTCTTCTCTGAAACCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 232
QY 1441 CAAAAGGCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1500
Db 231 CAAAAGGCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 172
QY 1501 TTGTGATTTCTTAAAGAAATGACTATATAATTTTATTTCCACTAAAAATATTTGTTCGC 1560
Db 171 TTGTGATTTCTTAAAGAAATGACTATATAATTTTATTTCCACTAAAAATATTTGTTCGC 112
QY 1561 ATTCAATTTTATAGCAACCAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1620
Db 111 ATTCAATTTTATAGCAACCAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 152
QY 1621 GCAAAATATGTTTAAATATAAAATGAAATTTGATTTTATAAAAAA 1671
Db 51 GCAAAATATGTTTAAATATAAAATGAAATTTGATTTATAAAAAA 1

RESULT 11
AR112510
LOCUS AR112510 1997 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 7 from patent US 6130071.
ACCESSION AR112510
VERSION AR112510.1 GI:14092410
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1997)
AUTHORS Alitalo, K. and Joukov, V.
TITLE Vascular endothelial growth factor C (VEGF-C) .DELTA. Cys.sub.156
JOURNAL Patent: US 6130071-A 7 10-OCT-2000;
FEATURES
source
1..1997
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 98.7%; Score 1651.8; DB 6; Length 1997;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTCTCTCCACCATGCATCTCGCTGGGCTTCTTCTCTGTGGGCTGTCTGTCTGCCCGCTG 60
Db 341 GTCTCTCCACCATGCATCTGTCTGGGCTTCTTCTCTGTGGGCTGTCTGTCTGCCCGCTG 400
QY 61 CGTGTCTCCCGGTCCTCGAGAGCGCCCGCGCGCGCTTCGAGTCGGACTCG 120
Db 401 CGTGTCTCCCGGTCCTCGAGAGCGCCCGCGCGCGCTTCGAGTCGGACTCG 460
QY 121 ACCTCTCGGACGGGAGCCCGACGCGGGGAGGCCACGGCTTATGCAAGCAAGATCTCG 180
Db 461 ACCTCTCGGACGGGAGCCCGACGCGGGGAGGCCACGGCTTATGCAAGCAAGATCTCG 520
```


QY 181 AGGACAGTACGGTCTGTGTCCAGTGTAGATGAACATCATGCTGTACTTACCCAGAT 240
DB 521 AGGACAGTACGGTCTGTGTCCAGTGTAGATGAACATCATGCTGTACTTACCCAGAT 580
QY 241 ATTGAAAAATGTACAAAGTGTACAGTAAGAAAGAGGCTGGCAACATACAGAAACAGG 300
DB 581 ATTGAAAAATGTACAAAGTGTACAGTAAGAAAGAGGCTGGCAACATACAGAAACAGG 640
QY 301 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGCAGCAATTAATACAG 360
DB 641 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGCAGCAATTAATACAG 700
QY 361 AGATCTTGAAGATTTGATTAATGATGGAGAGAGACTCAATGATGCGCAGGAGGTGT 420
DB 701 AGATCTTGAAGATTTGATTAATGATGGAGAGAGACTCAATGATGCGCAGGAGGTGT 760
QY 421 GTATAGATGTGGGAAAGAGTTTGGAGTCGGCAACAAACACCTTCTTTAAACCTCCATGTG 480
DB 761 GTATAGATGTGGGAAAGAGTTTGGAGTCGGCAACAAACACCTTCTTTAAACCTCCATGTG 820
QY 481 TGTCCGCTACAGATGTGGGAGTGTCTGCAATAGTAGGGGCTGCAGTGCATGAACACCA 540
DB 821 TGTCCGCTACAGATGTGGGAGTGTCTGCAATAGTAGGGGCTGCAGTGCATGAACACCA 880
QY 541 GCACGAGTACTCTCAGCAAGAGCTTATTTGAAATTTACAGTGCCTCTCTCTCAAGGCCCA 600
DB 881 GCACGAGTACTCTCAGCAAGAGCTTATTTGAAATTTACAGTGCCTCTCTCTCAAGGCCCA 940
QY 601 AACCGTAAACATCAATTTTCCCAATCACATCTCTGCGGATGATGTCTAAACCTGGATG 660
DB 941 AACCGTAAACATCAATTTTCCCAATCACATCTCTGCGGATGATGTCTAAACCTGGATG 1000
QY 661 TTTACAGCAAGTTCATTTCCATTTATAGACGTTCCTGCCAGCAACACTACCAAGTGC 720
DB 1001 TTTACAGCAAGTTCATTTCCATTTATAGACGTTCCTGCCAGCAACACTACCAAGTGC 1060
QY 721 AGGACGCAACAGCACTGCCCAATACATTTGAAATTTACAGTGCCTCTCTCTCAAGGCCCA 780
DB 1061 AGGACGCAACAGCACTGCCCAATACATTTGAAATTTACAGTGCCTCTCTCTCAAGGCCCA 1120
QY 781 TGGCTCAGGAAGATTTATGTTTCTCCGATGTGAGATGATCAACAGATGATTC 840
DB 1121 TGGCTCAGGAAGATTTATGTTTCTCCGATGTGAGATGATCAACAGATGATTC 1180
QY 841 ATGACATCTGTGGACCAACAGAGCTGGATGAAGAGACCTGTCACTGTCTGACAGAG 900
DB 1181 ATGACATCTGTGGACCAACAGAGCTGGATGAAGAGACCTGTCACTGTCTGACAGAG 1240
QY 901 CGGGCTTCGGCTGCCAGTGTGGACCCCAACAAAGAACTAGACAGAAACTCATGCGAGT 960
DB 1241 CGGGCTTCGGCTGCCAGTGTGGACCCCAACAAAGAACTAGACAGAAACTCATGCGAGT 1300
QY 961 GTGTCTGTAAAAACAACTCTTCCCGCAATGTGGGCAACCGAGAAATTTGATGAAA 1020
DB 1301 GTGTCTGTAAAAACAACTCTTCCCGCAATGTGGGCAACCGAGAAATTTGATGAAA 1360
QY 1021 ACACATGCCAGTGTGTATGTAAAAAGAACTGCCCCAGAAATCAACCCCTAAATCTGGAA 1080
DB 1361 ACACATGCCAGTGTGTATGTAAAAAGAACTGCCCCAGAAATCAACCCCTAAATCTGGAA 1420
QY 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAAAGAAAGTTC 1140
DB 1421 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAAAGAAAGTTC 1480
QY 1141 ACCACCAACATGACAGTGTGTACAGAGGCGCATGTAGCAACCGCAGAGGCTTGTGAGC 1200
DB 1481 ACCACCAACATGACAGTGTGTACAGAGGCGCATGTAGCAACCGCAGAGGCTTGTGAGC 1540
QY 1201 CAGGATTTTCATATAGTGAAGAGTGTGTGTTGTGCTCCCTTCATATATGCAAGACCA 1260
DB 1541 CAGGATTTTCATATAGTGAAGAGTGTGTGTTGTGCTCCCTTCATATATGCAAGACCA 1600
QY 1261 AAATGAGCTAAGATTTGTACTGTTTCCAGTTTCATCGATTTTCTATTATGAAACCTGTGT 1320

DB 1601 AAATGAGCTAAGATTTGTACTGTTTCCAGTTCATCGATTTTCTATTATGAAAACTGTGT 1660
QY 1321 TGCACAGTAGAAGTCTGTGTGAACAGAGAGACCTTTGGGTCCATGCTAACAAAGACA 1380
DB 1661 TGCACAGTAGAAGTCTGTGTGAACAGAGAGACCTTTGGGTCCATGCTAACAAAGACA 1720
QY 1381 AAAGTCTGTCTTTCTGAAACCACTGTGGATAACTTTACAGAAATGGAAGTCAATCTG 1440
DB 1721 AAAGTCTGTCTTTCTGAAACCACTGTGGATAACTTTACAGAAATGGAAGTCAATCTG 1780
QY 1441 CAAAAGCCCTCTGTAAAGACTGGTTTTCTGCCAATGACCAACAGCCAGATTTTCCTC 1500
DB 1781 CAAAAGCCCTCTGTAAAGACTGGTTTTCTGCCAATGACCAACAGCCAGATTTTCCTC 1840
QY 1501 TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAAAATATTGTTCTGC 1560
DB 1841 TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAAAATATTGTTCTGC 1900
QY 1561 ATTCAATTTATAGCAACAAATTTGGTAAACTCACTGTGATCAATATTTTATATCAT 1620
DB 1901 ATTCAATTTATAGCAACAAATTTGGTAAACTCACTGTGATCAATATTTTATATCAT 1960
QY 1621 GCAAAATATGTTTAAATAAATAAATGAAATTTGTAAT 1655
DB 1961 GCAAAATATGTTTAAATAAATAAATGAAATTTGTAAT 1995

RESULT 12
ARI57589
LOCUS
DEFINITION Sequence 44 from patent US 6245530.
ACCESSION ARI57589
VERSION ARI57589.1 GI:16218541
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1997)
AUTHORS Alitalo, K. and Joukov, V.
TITLE Receptor ligand
JOURNAL Patent: US 6245530-A 44 12-JUN-2001;
FEATURES
Location/Qualifiers
1..1997
source
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 98.7%; Score 1651.8; DB 6; Length 1997;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GTCCCTCCACCATGCACTGCTGGGCTTCTCTCTGTGGCGTGTCTCTGCTCGCGGTG 60
DB 341 GTCCCTCCACCATGCACTGCTGGGCTTCTCTCTGTGGCGTGTCTCTGCTCGCGGTG 400
QY 61 CGCTGCTCCCGGCTCTCTCGCAGGGCGCGCGCGCGCGCGCGCTTCGAGTCCGAGCTCG 120
DB 401 CGCTGCTCCCGGCTCTCTCGCAGGGCGCGCGCGCGCGCGCTTCGAGTCCGAGCTCG 460
QY 121 ACCTCTCGAGCCGAGGCCCGCAGCGGGCGAGGCCACCGGCTTATGCAAGCAAGATCTGG 180
DB 461 ACCTCTCGAGCCGAGGCCCGCAGCGGGCGAGGCCACCGGCTTATGCAAGCAAGATCTGG 520
QY 181 AGGAGCAGTTACGGTCTGTGTCCAGTGTAGTGAACATCATGCTGTACTCTTACCAGAA 240
DB 521 AGGAGCAGTTACGGTCTGTGTCCAGTGTAGTGAACATCATGCTGTACTCTTACCAGAA 580
QY 241 ATTGGAATGTACAAAGTGTACAGTAAGAAAGAGGCTGGCAACATTAACAGAAACAGG 300
DB 581 ATTGGAATGTACAAAGTGTACAGTAAGAAAGAGGCTGGCAACATTAACAGAAACAGG 640
QY 301 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGCAGCAATTAATACAG 360

Db 641 CCAACCTCACTCAAGGACAGAGACTATAAATTTGCTGACGACATTAATAACAG 700
Qy 361 AGATCTTGAAAGATTGATAATAGTGGAGAAAGACTCAATGCATGCCACGGAGGTGT 420
Db 701 AGATCTTGAAAGATTGATAATAGTGGAGAAAGACTCAATGCATGCCACGGAGGTGT 760
Qy 421 GTATAGATGGGGAAGAGTTTGGAGTCCGCGAACAACACCTCTTTAAACCTCCATGTG 480
Db 761 GTATAGATGGGGAAGAGTTTGGAGTCCGCGAACAACACCTCTTTAAACCTCCATGTG 820
Qy 481 TGTCCGCTACAGATGTGGGGTTGCTGCAATAGTGGGGGCTGCAGTGCATGAACACCA 540
Db 821 TGTCCGCTACAGATGTGGGGTTGCTGCAATAGTGGGGGCTGCAGTGCATGAACACCA 880
Qy 541 GCAGAGTACTCAGCAAGACGTATTGAAATTACAGTGCCTCTCTCTCAAGGCCCA 600
Db 881 GCAGAGTACTCAGCAAGACGTATTGAAATTACAGTGCCTCTCTCTCAAGGCCCA 940
Qy 601 AACCAATAAATCAGTTTGGCAATCACCTTCTCGCATGCATGTCTAAACTGGATG 660
Db 941 AACCAATAAATCAGTTTGGCAATCACCTTCTCGCATGCATGTCTAAACTGGATG 1000
Qy 661 TTTACAGACAGTTTCATTCATTAATAGAGTTCCTCGCAGCAAACTACCAACAGTGC 720
Db 1001 TTTACAGACAGTTTCATTCATTAATAGAGTTCCTCGCAGCAAACTACCAACAGTGC 1060
Qy 721 AGGAGGGAACAAGACCTGCCCAACCAATTACATGTGGAATATCAATCTGCAGATGC 780
Db 1061 AGGAGGGAACAAGACCTGCCCAACCAATTACATGTGGAATATCAATCTGCAGATGC 1120
Qy 781 TGGCTCAGGAAGATTTATGTTTTCCTCGGATGTGGAGTGAATCAACAGATGGATTC 840
Db 1121 TGGCTCAGGAAGATTTATGTTTTCCTCGGATGTGGAGTGAATCAACAGATGGATTC 1180
Qy 841 ATGATCTGTGGACCAACAAGAGCTGGATGAGAGACCTGTCAGTGTCTGCAGAG 900
Db 1181 ATGATCTGTGGACCAACAAGAGCTGGATGAGAGACCTGTCAGTGTCTGCAGAG 1240
Qy 901 CGGGGCTTCGGCTCGCAGCTGTGGACCCCAAGAACTAGACAGAACTCATGCGCAGT 960
Db 1241 CGGGGCTTCGGCTCGCAGCTGTGGACCCCAAGAACTAGACAGAACTCATGCGCAGT 1300
Qy 961 GTGCTGTAAAAACAATCTCTCCAGGCAATGTGGGGCAACCGAGAAATTTGATGAAA 1020
Db 1301 GTGCTGTAAAAACAATCTCTCCAGGCAATGTGGGGCAACCGAGAAATTTGATGAAA 1360
Qy 1021 ACATATGCCAGTGTATGTAAABAGACCTGCCCGCAGAAATCAACCCCTAAATCCTGGA 1080
Db 1361 ACATATGCCAGTGTATGTAAABAGACCTGCCCGCAGAAATCAACCCCTAAATCCTGGA 1420
Qy 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTGTTAAAGGAAAGAAAGTTCC 1140
Db 1421 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTGTTAAAGGAAAGAAAGTTCC 1480
Qy 1141 ACCACCAACAATGACAGCTGTACAGCGGCATGTAGAACCGCAGAGAGCTTGTGAGC 1200
Db 1481 ACCACCAACAATGACAGCTGTACAGCGGCATGTAGAACCGCAGAGAGCTTGTGAGC 1540
Qy 1201 CAGGATTTTCATATAGTGAAGAAAGTGTGCTGTGTCCTTCATATTTGCAAGAACCCAC 1260
Db 1541 CAGGATTTTCATATAGTGAAGAAAGTGTGCTGTGTCCTTCATATTTGCAAGAACCCAC 1600
Qy 1261 AAATGAGCTAAGATTGACTGTTTCCAGTTCATCGATTTTCTATTATGAAACCTGTGT 1320
Db 1601 AAATGAGCTAAGATTGACTGTTTCCAGTTCATCGATTTTCTATTATGAAACCTGTGT 1660
Qy 1321 TGCCACAGTAGAATGTCTGTGAACAGAGAGACCTTTGTGGGTCATGTGTACAAAGACA 1380
Db 1661 TGCCACAGTAGAATGTCTGTGAACAGAGAGACCTTTGTGGGTCATGTGTGTACAAAGACA 1720
Qy 1381 AAATGCTGTCTTCTCAACCAATGTGGATAACTTTTACAGAAATGGACTGAGCTCATCTG 1440

Db 1721 AAGTCTGTCTTCTCTGAACCATGTGGATAACTTTTACAGAAATGAGCTGGAGCTCATCTG 1780
Qy 1441 CAAAAGGGCTCTTGTAAAGAGCTGTTTTCTGCAATGACCAACAGCAAGATTTTCTCTC 1500
Db 1781 CAAAAGGGCTCTTGTAAAGAGCTGTTTTCTGCAATGACCAACAGCAAGATTTTCTCTC 1840
Qy 1501 TTGTGATTTCTTTAAAGAAAGTACTATATAATTTATTTTCCACTAAAAATATTTTCTGC 1560
Db 1841 TTGTGATTTCTTTAAAGAAAGTACTATATAATTTATTTTCCACTAAAAATATTTTCTGC 1900
Qy 1561 ATTCAATTTTATAGCAACAACAATTTGGTAAACCTCACTGTGATCAATATTTTATATCAT 1620
Db 1901 ATTCAATTTTATAGCAACAACAATTTGGTAAACCTCACTGTGATCAATATTTTATATCAT 1960
Qy 1621 GCAAAATATGTTTAAATAAAAATGAAATTTGTATT 1655
Db 1961 GCAAAATATGTTTAAATAAAAATGAAATTTGTATT 1995

RESULT 13
BD247178 1997 bp DNA linear PAT 17-JUL-2003
DEFINITION Use of VEGF-C or VEGF-D gene or protein to prevent restenosis.
ACCESSION BD247178
VERSION BD247178.1 GI:33056948
KEYWORDS JP 2002528420-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1997)
AUTHORS Hertzuala,S.Y., Alitalo,K., Hiltunen,M.O., Jeltsch,M.M. and Achen,M.G.
TITLE Use of VEGF-C or VEGF-D gene or protein to prevent restenosis
JOURNAL Patent: JP 2002528420-A 1 03-SEP-2002;
LUDWIG INSTITUTE FOR CANCER RESEARCH, HELSINKI UNIVERSITY LICENSING
LTD OY, SEPO YLA HERTTUALA
COMMENT OS Homo sapiens (human)
PN JP 2002528420-A/1
PD 03-SEP-2002
PF 26-OCT-1999 JP 2000578021
PR 28-OCT-1998 US 60/105587
PI SEPO YLA HERTTUALA, KARI ALITALO, MIKKO O HILTUNEN, MARKKU M PI
JELTSCH,
PI MARC G ACHEN
PC A61K38/22,A61K31/711,A61K48/00,A61L33/00,A61L33/00,A61M25/00,
PC A61M29/02,A61P41/00,A61P43/00,C07K14/475,C12N15/09,A61K37/24,
PC A61P7/02,A61P41/00,A61P43/00,C07K14/475,C12N15/09,A61K37/24,
PC C12N15/00,
PC A61L33/00,A61L33/00
CC Use of VEGF-C or VEGF-D gene or protein to prevent restenosis
FH Key Location/Qualifiers
FT CDS (352)..(1608).
FEATURES
source
1..1997
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 98.7%; Score 1651.8; DB 6; Length 1997;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 GTCCTTCACCATGACACTCGCTGGGCTTCTTCTCTGTGGGTGTCTCTGTCTGCCCTG 60
Db 341 GTCCTTCACCATGACACTCGCTGGGCTTCTTCTCTGTGGGTGTCTCTGTCTGCCCTG 400
Qy 61 CGCTGCTCCCGGTCTCTCGAGCGCCGCCGCCGCCCTTCAGTCCGAGATCG 120
Db 401 CGCTGCTCCCGGTCTCTCGAGCGCCGCCGCCGCCCTTCAGTCCGAGATCG 460
Qy 121 ACCTCTCGAGCGGAGCCCGACGCGGCGGAGGCCACGGCTTATGCAAGCAAGATCTGG 180

461 ACCTCTCGAGCGCGAGCCGACCGCGCGAGGCCACGGCTTATGCAAGCAAGATCTGG 520
181 AGGAGCAGTTACGGCTGTGTCCAGTGTAGATGAACCTCATGACTGTACTCTACCCAGAAT 240
521 AGGAGCAGTTACGGCTGTGTCCAGTGTAGATGAACCTCATGACTGTACTCTACCCAGAAT 580
241 ATTGGAATAATGTACAAGTGTCAAGTAAAGGAAGAGGCTGGCAACATAACAGAGAAGCAGG 300
581 ATTGGAATAATGTACAAGTGTCAAGTAAAGGAAGAGGCTGGCAACATAACAGAGAAGCAGG 640
301 CCACCTCAACTCAAGACAGAGAGACTATAAATTTGCTGAGAGCAATTAATACAG 360
641 CCACCTCAACTCAAGACAGAGAGACTATAAATTTGCTGAGAGCAATTAATACAG 700
361 AGATCTTTGAAAGTATTGATAATCAGTGGAGAAAGACTCAATGCAATGCCAGGGAGGTGT 420
701 AGATCTTTGAAAGTATTGATAATCAGTGGAGAAAGACTCAATGCAATGCCAGGGAGGTGT 760
421 GTATAGATGTGGGAAGAGGTTTGGAGTGGGACAAACACCTCTTTAAACCTCCATGTG 480
761 GTATAGATGTGGGAAGAGGTTTGGAGTGGGACAAACACCTCTTTAAACCTCCATGTG 820
481 TGTCCGCTCAGAGATGTGGGGTTGCTGCAATAGTGAAGGGCTGCAATGATGAACACCA 540
821 TGTCCGCTCAGAGATGTGGGGTTGCTGCAATAGTGAAGGGCTGCAATGATGAACACCA 880
541 GCACGAGTACTCTCAGCAAGACGTTATTGAAATTAAGTGCCTCTCTCTCAAGGGCCCCA 600
881 GCACGAGTACTCTCAGCAAGACGTTATTGAAATTAAGTGCCTCTCTCTCAAGGGCCCCA 940
601 AACCACTAACCAATCAGTTTCCCAATCACCTTCCGATGATGATCTCTAACTGGATG 660
941 AACCACTAACCAATCAGTTTCCCAATCACCTTCCGATGATGATCTCTAACTGGATG 1000
661 TTTACAGCAAGTTCATTCATTAATTAGACGTTCCCTGCCAGCAACACTACCAAGGTGC 720
1001 TTTACAGCAAGTTCATTCATTAATTAGACGTTCCCTGCCAGCAACACTACCAAGGTGC 1060
721 AGGAGCGCAACAGACCTGCCCAATACATGATGGAATTAATCAATCTGCAGATGCC 780
1061 AGGAGCGCAACAGACCTGCCCAATACATGATGGAATTAATCAATCTGCAGATGCC 1120
781 TGGCTCAGCAAGATTTATGTTTCTCCGATGCTGGAGTGAATCAACAGATGGATCC 840
1121 TGGCTCAGCAAGATTTATGTTTCTCCGATGCTGGAGTGAATCAACAGATGGATCC 1180
841 ATGACATCTGGACCAACAGAGCTGGATGAAGAGACCTGTCACTGTGTCTGCAGAG 900
1181 ATGACATCTGGACCAACAGAGCTGGATGAAGAGACCTGTCACTGTGTCTGCAGAG 1240
901 CGGGCTTCGGCTGCCAGCTGTGACCCCAAGAGACTAGACAGAACTCATGCCAGT 960
1241 CGGGCTTCGGCTGCCAGCTGTGACCCCAAGAGACTAGACAGAACTCATGCCAGT 1300
961 GTGTCTGTAAACAACTCTTCCAGCAATGTGGGGCCCAACCGAGAAATTTGATGAAA 1020
1301 GTGTCTGTAAACAACTCTTCCAGCAATGTGGGGCCCAACCGAGAAATTTGATGAAA 1360
1021 ACACATGCCAGTGTGTATGAAGAACTGCCCCCAGAAATCAACCCCTAATCTCTGAA 1080
1361 ACACATGCCAGTGTGTATGAAGAACTGCCCCCAGAAATCAACCCCTAATCTCTGAA 1420
1081 AATGTCCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGAAAGAAAGTTCC 1140
1421 AATGTCCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGAAAGAAAGTTCC 1480
1141 ACACCAACATGAGCTGTATACAGAGGCTGTACGAAACCGCCAGAGGCTGTGAGC 1200
1481 ACCACCAACATGAGCTGTATACAGAGGCTGTACGAAACCGCCAGAGGCTGTGAGC 1540
1201 CAGGATTTTCAATATAGTGAAGAGTGTGTGTTGTGCTTCTCATATTTGGCAAGACCCAC 1260

1541 CAGGATTTTCAATATAGTGAAGAGTGTGTGTTGTTCCCTTCTCATATTTGGAAGACCCAC 1600
1261 AAATCAGCTAAGATTTGTACTGTTTTCCAGTTTCATGATTTTCTATATATGGAAGACTGTG 1320
1601 AAATCAGCTAAGATTTGTACTGTTTTCCAGTTTCATGATTTTCTATATATGGAAGACTGTG 1660
1321 TGCCACAGTAGAAGTGTCTGTGAACAGAGAGACCTTGTGGTCCATCTCAACAAGACA 1380
1661 TGCCACAGTAGAAGTGTCTGTGAACAGAGAGACCTTGTGGTCCATCTCAACAAGACA 1720
1381 AAAGTCTGTCTTTCTCTGAAACCATGTGGATAACTTTTACAGAAATGGAGCTCATCTG 1440
1721 AAAGTCTGTCTTTCTCTGAAACCATGTGGATAACTTTTACAGAAATGGAGCTCATCTG 1780
1441 CAAAAGGCTCTTTGTAAGAAGTGTGTTTCTGCCAATGACCAACAGCCAGATTTTCTC 1500
1781 CAAAAGGCTCTTTGTAAGAAGTGTGTTTCTGCCAATGACCAACAGCCAGATTTTCTC 1840
1501 TTGTGATTTCTTTTAAAGAAAGTACTATATAATTTATTTCCACTAAATAATTTGTTCTGC 1560
1841 TTGTGATTTCTTTTAAAGAAAGTACTATATAATTTATTTCCACTAAATAATTTGTTCTGC 1900
1561 ATTCAATTTTATAGCAACAAATTTGTAAGAAGTCACTGTGATCAATATTTTATATCAT 1620
1901 ATTCAATTTTATAGCAACAAATTTGTAAGAAGTCACTGTGATCAATATTTTATATCAT 1960
1621 GCATAATATGTTTAAATAAATAAATGAATTTGTTATT 1655
1961 GCATAATATGTTTAAATAAATAAATGAATTTGTTATT 1995

RESULT 14
AR201986
LOCUS AR201986 1997 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 7 from patent US 6361946.
ACCESSION AR201986
VERSION AR201986.1 GI:20256525
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1997)
AUTHORS Alitalo, K. and Joukov, V.
TITLE Vascular endothelial growth factor C (VEGF-C) protein and gene, mutants thereof, and uses thereof
JOURNAL Patent: US 6361946-A 7 26-MAR-2002;
FEATURES
Location/Qualifiers
source 1..1997
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 98.7%; Score 1651.8; DB 6; Length 1997;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
1 GTCTTCCACCAATGCACTCGTGGCTTCTTCTGTGGCGTGTCTCTGCTCGCGCTG 60
341 GTCTTCCACCAATGCACTTGTGGCTTCTTCTGTGGCGTGTCTCTGCTCGCGCTG 400
61 CGCTGCTCCGGTCTCTCGAGGGCCCGCGCGCGCGCGCTTCGAGTCCGAGTCTG 120
401 CGCTGCTCCGGTCTCTCGAGGGCCCGCGCGCGCGCGCTTCGAGTCCGAGTCTG 460
121 ACCTCTCGAGCGCGAGCCCGAGCGCGCGCGCGCTTCGAGTCCGAGTCTG 180
461 ACCTCTCGAGCGCGAGCCCGAGCGCGCGCGCGCTTCGAGTCCGAGTCTG 520
181 AGAGCAGTTACGCTGTGTCCAGTGTAGATGAATCTATGCTACTCTACCCAGAAT 240
521 AGAGCAGTTACGCTGTGTCCAGTGTAGATGAATCTATGCTACTCTACCCAGAAT 580
241 ATTGGAATAATGTACAAGTGTCACTAAGGAAGAGGCTGGCAACATAACAGAGAAGCAGG 300

Db 581 ATTTGGAATATGTAAGTGTAGCTAGGTAAGGAGGAGGCTGGCAACATAACAGAGAACAGG 640
Qy 301 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGAGACACATTAATAACAG 360
Db 641 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGAGACACATTAATAACAG 700
Qy 361 AGATCTTTGAAAGTATTGATTAATAGTGGAGAAAGACTCAATGATGCCACGGAGGTGT 420
Db 701 AGATCTTTGAAAGTATTGATTAATAGTGGAGAAAGACTCAATGATGCCACGGAGGTGT 760
Qy 421 GTATAGATGTGGGAGAGGTTTGGAGTCCGACAAACACCTCTTTAAACCTCCATGTG 480
Db 761 GTATAGATGTGGGAGAGGTTTGGAGTCCGACAAACACCTCTTTAAACCTCCATGTG 820
Qy 481 TGTCCGCTCTACAGATGTGGGGTGTCTGCAATAGTGGGGGCTGCAGTGCAATGAAACCA 540
Db 821 TGTCCGCTCTACAGATGTGGGGTGTCTGCAATAGTGGGGGCTGCAGTGCAATGAAACCA 880
Qy 541 GCAGGAGTACCTCAGCAAGAGCTTATTGAAATTTAGAGTGCCTCTCTCAAGGCCCA 600
Db 881 GCAGGAGTACCTCAGCAAGAGCTTATTGAAATTTAGAGTGCCTCTCTCAAGGCCCA 940
Qy 601 AACAGTAAACAATCAGTCTTTGCCAATCACACTTCTTCCGATGATGTCTTAACTGGATG 660
Db 941 AACAGTAAACAATCAGTCTTTGCCAATCACACTTCTTCCGATGATGTCTTAACTGGATG 1000
Qy 661 TTTACAGCAAGTTCATTCATTTATGTTTCTCGGATGTGGAGATGATCAACAGATGATTC 720
Db 1001 TTTACAGCAAGTTCATTCATTTATGTTTCTCGGATGTGGAGATGATCAACAGATGATTC 1060
Qy 721 AGGAGCGAACAAGACCTGCCCCACCAATTTACATGTGGAATATCACTCTGAGATGCC 780
Db 1061 AGGAGCGAACAAGACCTGCCCCACCAATTTACATGTGGAATATCACTCTGAGATGCC 1120
Qy 781 TGGCTCAGGAAGATTTATGTTTCTCGGATGTGGAGATGATCAACAGATGATTC 840
Db 1121 TGGCTCAGGAAGATTTATGTTTCTCGGATGTGGAGATGATCAACAGATGATTC 1180
Qy 841 ATGACATCTGGGACCAACAAGAGAGCTGGATGAAGAGCTCTGAGTGTCTGAGAG 900
Db 1181 ATGACATCTGGGACCAACAAGAGAGCTGGATGAAGAGCTCTGAGTGTCTGAGAG 1240
Qy 901 CGGGGCTTGGGCTTCCGAGCTGTGGACCCCAAGAACTAGACAGAACTCATGCCAGT 960
Db 1241 CGGGGCTTGGGCTTCCGAGCTGTGGACCCCAAGAACTAGACAGAACTCATGCCAGT 1300
Qy 961 GTGCTGTAAACAACTCTTCCAGCCCAATGTGGGGCCCAAGAGATTTGATGAA 1020
Db 1301 GTGCTGTAAACAACTCTTCCAGCCCAATGTGGGGCCCAAGAGATTTGATGAA 1360
Qy 1021 ACATGCGCAGTGTATGTAAAGAACTTCCGAGAACTCAACCCCTAAATCCTGGAA 1080
Db 1361 ACATGCGCAGTGTATGTAAAGAACTTCCGAGAACTCAACCCCTAAATCCTGGAA 1420
Qy 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAGAGTTCC 1140
Db 1421 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAGAGTTCC 1480
Qy 1141 ACCACCAACATGAGCTGTTTACAGCGGCACTGTACGACCGCAGAGGCTTGTGAGC 1200
Db 1481 ACCACCAACATGAGCTGTTTACAGCGGCACTGTACGACCGCAGAGGCTTGTGAGC 1540
Qy 1201 CAGGATTTTCATATAGTGAAGAGTGTGCTGTTGTCCTTTCATATGCGAAAGACAC 1260
Db 1541 CAGGATTTTCATATAGTGAAGAGTGTGCTGTTGTCCTTTCATATGCGAAAGACAC 1600
Qy 1261 AAATGAGCTAGATTTGATCTTTTCCAGTTTCATGCTTTTCTATTATGAAACTGTGT 1320
Db 1601 AAATGAGCTAGATTTGATCTTTTCCAGTTTCATGCTTTTCTATTATGAAACTGTGT 1660
Qy 1321 TGCCACAGTACAGTCTGTGAAACAGAGAGACCTTGTGGGTCCATGCTTAACAAAGACA 1380

Db 1661 TGCCACAGTACAGTCTGTGTAACAGAGAGACCCCTTGTGGTCCATGCTAAACAAGACA 1720
Qy 1381 AAAGTCTGTCTTTCTTCTGAACCATGTGGATAAATTTACAGAAATGAGACTGGAGCTCATCTG 1440
Db 1721 AAAGTCTGTCTTTCTTCTGAACCATGTGGATAAATTTACAGAAATGAGACTGGAGCTCATCTG 1780
Qy 1441 CAAAGGCTCTTCTTGAAGACTGGTCTTCTGCCAATGACCAACAGCAAGATTTTCTC 1500
Db 1781 CAAAGGCTCTTCTTGAAGACTGGTCTTCTGCCAATGACCAACAGCAAGATTTTCTC 1840
Qy 1501 TTGTGATTTCTTTTAAAGAAATGACTATATAATTTATTTCCACTAAAAATATTTCTTC 1560
Db 1841 TTGTGATTTCTTTTAAAGAAATGACTATATAATTTATTTCCACTAAAAATATTTCTTC 1900
Qy 1561 ATTCAATTTTATAGCAACACAAATTTGTTAAACTCACTGTGATCAATATTTTATATCAT 1620
Db 1901 ATTCAATTTTATAGCAACACAAATTTGTTAAACTCACTGTGATCAATATTTTATATCAT 1960
Qy 1621 GCAAAATATGTTTAAATATAAATGAAAATTTGTAT 1655
Db 1961 GCAAAATATGTTTAAATATAAATGAAAATTTGTAT 1995

RESULT 15
AR212635
LOCUS 1997 bp DNA linear PAT 25-SEP-2002
DEFINITION Sequence 32 from patent US 6403088.
ACCESSION AR212635
VERSION AR212635.1 GI:23309413
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1997)
AUTHORS Alitalo, K. and Joukov, V.
TITLES Antibodies reactive with VEGF-C, a ligand for the Flt-4 receptor
JOURNAL tyrosine kinase (VEGFR-3)
PUBLISHED Patent: US 6403088-A 32 11-JUN-2002;
FEATURES Location/Qualifiers
1..1997
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 98.7%; Score 1651.8; DB 6; Length 1997;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 GTCTCTCCACCATGCACTCGCTGGGCTTCTTCTGTGGCGTGTCTCTGCTCGCGCTG 60
Db 341 GTCTCTCCACCATGCACTCGCTGGGCTTCTTCTGTGGCGTGTCTCTGCTCGCGCTG 400
Qy 61 CGCTCTCCCGGCTCTCCGAGGCGCCGCGCGCGCGCGCTTCGAGTCCGAGCTCG 120
Db 401 CGCTCTCCCGGCTCTCCGAGGCGCCGCGCGCGCGCTTCGAGTCCGAGCTCG 460
Qy 121 ACCTCTCGGACCGGAGCCGACGCGGCGGCGGCGGCTTATGCAAGCAAGATCTGG 180
Db 461 ACCTCTCGGACCGGAGCCGACGCGGCGGCGGCGGCTTATGCAAGCAAGATCTGG 520
Qy 181 AGGAGCAGTTACGGTCTGTGTCTGCTAGATGAATCATGCTGTACTTACCCAGAT 240
Db 521 AGGAGCAGTTACGGTCTGTGTCTGCTAGATGAATCATGCTGTACTTACCCAGAT 580
Qy 241 ATTGGAATATGTAAGTGTGCTAGTAAGAAAGGCGGCGGCGGCTTATGCAAGCAAGATCTGG 300
Db 581 ATTGGAATATGTAAGTGTGCTAGTAAGAAAGGCGGCGGCGGCTTATGCAAGCAAGATCTGG 640
Qy 301 CCAACCTCAACTCAAGGACAGAGAGCTATAAAATTTGCTGCGACCATTAATATACAG 360
Db 641 CCAACCTCAACTCAAGGACAGAGAGCTATAAAATTTGCTGCGACCATTAATATACAG 700
Qy 361 AGATCTTGAAGATATTGATTAATAGTGGAGAAAGACTCAATGATGCCACGGAGGTGT 420

Db	701	AGATCTTTGAAAGTATTGATTAATGAGTGGAGAAAGACTCAATGTCATGCCGAGGAGTGT	760
Qy	421	GTATAGATGTGGGAAGAGTTTGGAGTCGGGCAACACACCTCTTTTAAACCTCCATGTG	480
Db	761	GTATAGATGTGGGAAGAGTTTGGAGTCGGGCAACACACCTCTTTTAAACCTCCATGTG	820
Qy	481	TGTCGGTCTACAGATGTGGGGTTTCTGTGCAATAGTGAAGGGCTGCAGTGAATGAACCA	540
Db	821	TGTCGGTCTACAGATGTGGGGTTTCTGTGCAATAGTGAAGGGCTGCAGTGAATGAACCA	880
Qy	541	GCACGAGTCTCAGCAAGACGTTATTTGAAATTACAGTGCCTCTCTCTCAGGCCCCA	600
Db	881	GCACGAGTCTCAGCAAGACGTTATTTGAAATTACAGTGCCTCTCTCTCAGGCCCCA	940
Qy	601	AACCAAGTAAACATCAGTTTTCGCAATCACACTTCTGCGCATGTCCTTAACTGGATG	660
Db	941	AACCAAGTAAACATCAGTTTTCGCAATCACACTTCTGCGCATGTCCTTAACTGGATG	1000
Qy	661	TTTACAGACAAGTTCAITTCATTTATAGACGTTCCCTGCCAGCAACACTACCAGTGC	720
Db	1001	TTTACAGACAAGTTCAITTCATTTATAGACGTTCCCTGCCAGCAACACTACCAGTGC	1060
Qy	721	AGGACGCAACAGACCTGCCCCCAATACATGTGGAATAACATCATCTGCAGATGCC	780
Db	1061	AGGACGCAACAGACCTGCCCCCAATACATGTGGAATAACATCATCTGCAGATGCC	1120
Qy	781	TGGCTCAGAGAATTTATGTTTTCCTCGGATGCTGGAGATGACTCAACAGATGGATTC	840
Db	1121	TGGCTCAGAGAATTTATGTTTTCCTCGGATGCTGGAGATGACTCAACAGATGGATTC	1180
Qy	841	ATGACATCTGTGGACCAACAGAGCTGGATGAAGACCTGTGAGTGTCTGTCAGAG	900
Db	1181	ATGACATCTGTGGACCAACAGAGCTGGATGAAGACCTGTGAGTGTCTGTCAGAG	1240
Qy	901	CGGGCTTCGGCTGCCAGCTGTGGAACCCCAAGAACTAGACAGAACTCATGCCAGT	960
Db	1241	CGGGCTTCGGCTGCCAGCTGTGGAACCCCAAGAACTAGACAGAACTCATGCCAGT	1300
Qy	961	GTGTCGTGTAACAACTCTTCCCGACGCAATGTGGGGCCCAACGGAAATTTGATGAA	1020
Db	1301	GTGTCGTGTAACAACTCTTCCCGACGCAATGTGGGGCCCAACGGAAATTTGATGAA	1360
Qy	1021	ACACATGCCAGTGTATGTAAGAACTGTGCCCGAGAAATCAACCCCTTAATCCTGGAA	1080
Db	1361	ACACATGCCAGTGTATGTAAGAACTGTGCCCGAGAAATCAACCCCTTAATCCTGGAA	1420
Qy	1081	AATGTCCTGTGAATGTACAGAACTCCACAGAAATGCTTGTGTAAGGAAGAGTTC	1140
Db	1421	AATGTCCTGTGAATGTACAGAACTCCACAGAAATGCTTGTGTAAGGAAGAGTTC	1480
Qy	1141	ACCACCAACATGCAGCTGTACAGCGGCCCATGTACGAACCGCCAGAGGCTTGTGAGC	1200
Db	1481	ACCACCAACATGCAGCTGTACAGCGGCCCATGTACGAACCGCCAGAGGCTTGTGAGC	1540
Qy	1201	CAGGATTTTCATATAGTGAAGAGTGTGTCGTTGTGTCCTTCATATTGGCAAGACCAC	1260
Db	1541	CAGGATTTTCATATAGTGAAGAGTGTGTCGTTGTGTCCTTCATATTGGCAAGACCAC	1600
Qy	1261	AAATGAGCTAAGATTGTACTGTTTCCAGTTCATCGATTTTCTATTATGGAAGACTGT	1320
Db	1601	AAATGAGCTAAGATTGTACTGTTTCCAGTTCATCGATTTTCTATTATGGAAGACTGT	1660
Qy	1321	TGCCACAGTAGAATCTGTGTGAACAGAGAGACCTTGTGGGTCCATGCTAACCAAGACA	1380
Db	1661	TGCCACAGTAGAATCTGTGTGAACAGAGAGACCTTGTGGGTCCATGCTAACCAAGACA	1720
Qy	1381	AAAGTCTGCTTTCTGACCACTGTGATTAATTTACAGAAATGGAGCTGGAGCTCATCTG	1440
Db	1721	AAAGTCTGCTTTCTGACCACTGTGATTAATTTACAGAAATGGAGCTGGAGCTCATCTG	1780
Qy	1441	CAAAAGGCTCTGTGTAAGAGCTGTTTCTGCAATGACCAACAGCCAGATTTTCTCTC	1500

Search completed: February 25, 2004, 20:57:49
Job time : 6423 secs

Db	1781	CAAAAAGGCTCTTGTAAAGACGTGGTTTTCTGCCAATGACCAAAACAGCCAGATTTTCTCTC	1840
Qy	1501	TTGTGATTTTCTTTTAAAAAGAAATGACTATATAATTTATTTTCCACTAAAAATATTGTTCTGC	1560
Db	1841	TTGTGATTTTCTTTTAAAAAGAAATGACTATATAATTTATTTTCCACTAAAAATATTGTTCTGC	1900
Qy	1561	ATTCAATTTTATAGCAACAACTTGGTAAAACTCACTGTGATCAATATTTTATATCAT	1620
Db	1901	ATTCAATTTTATAGCAACAACTTGGTAAAACTCACTGTGATCAATATTTTATATCAT	1960
Qy	1621	GCAAAATATGTTTAAAAATAAAATGAAATTTGTATT	1655
Db	1961	GCAAAATATGTTTAAAAATAAAATGAAATTTGTATT	1995

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2004, 16:56:30 ; Search time 653 Seconds
(without alignments)
10890.480 Million cell updates/sec

Title: US-09-921-143-1
Perfect score: 1674
Sequence: 1 gtcttcacatgcactgc.....ttataaaaaaaaaaaaaa 1674

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1674	100.0	1674	2	AAX84837 Full leng
2	1674	100.0	1674	2	AAX10523 CDNA enco
3	1674	100.0	1674	3	AA52080 Vascular
4	1674	100.0	1674	4	AA91004 Human VEG
5	1674	100.0	1674	6	ABX10034 Human vas
6	1674	100.0	1674	7	ABT17108 Human VEG
7	1674	100.0	1674	7	ABQ76966 Human VEG
8	1674	100.0	1674	7	AA49522 Human vas
9	1667.6	99.6	1674	2	AAT51371 Human vas
10	1664.6	99.4	2031	2	AAT59929 Human vas
11	1651.8	98.7	1997	2	AAT84276 Human Flt
12	1651.8	98.7	1997	2	AAV52576 Human vas
13	1651.8	98.7	1997	3	AA000339 Human vas
14	1651.8	98.7	1997	3	AA62406 Human VEG
15	1651.8	98.7	1997	4	AA68953 Human VEG
16	1651.8	98.7	1997	4	AA312820 DNA enco
17	1651.8	98.7	1997	9	ADD08947 Human VEG
18	1651.8	98.7	2015	9	ADC64993 Human VEG
19	1651.8	98.7	2879	3	AA000353 CMV-VEGF-
20	1607.8	96.0	2321	2	AAT68811 Human fce
21	1536	91.2	1526	2	AAX84838 Truncated
22	1536	91.2	1526	2	AA10524 CDNA enco
23	1526	91.2	1526	4	AA91010 Human VEG

24	1526	91.2	1526	6	ABK10035 Human vas
25	1520.2	90.8	1525	3	AA52081 Truncated
26	1520.2	90.8	1525	7	ABT17109 Human VEG
27	1520.2	90.8	1525	7	ABQ76967 Human tru
28	1520.2	90.8	1525	7	AA49523 Human vas
29	1517	90.6	1525	2	AAT03950 DNA enco
30	1517	90.6	1525	6	ABK10062 Expressio
31	1517	90.6	1525	6	ABV78239 Human VEG
32	1517	90.6	1525	6	ABV78239 Human VEG
33	1517	90.6	1525	6	ABV78239 Human VEG
34	1517	90.6	1525	6	ABV78239 Human VEG
35	1517	90.6	1525	6	ABV78239 Human VEG
36	1517	90.6	1525	6	ABV78239 Human VEG
37	1517	90.6	1525	6	ABV78239 Human VEG
38	1517	90.6	1525	6	ABV78239 Human VEG
39	1517	90.6	1525	6	ABV78239 Human VEG
40	1517	90.6	1525	6	ABV78239 Human VEG
41	1517	90.6	1525	6	ABV78239 Human VEG
42	1517	90.6	1525	6	ABV78239 Human VEG
43	1517	90.6	1525	6	ABV78239 Human VEG
44	1517	90.6	1525	6	ABV78239 Human VEG
45	1517	90.6	1525	6	ABV78239 Human VEG

ALIGNMENTS

RESULT 1
AAX84837
ID AAX84837 standard; cDNA; 1674 BP.
XX
AC AAX84837;
XX
XX
DT 22-SEP-1999 (first entry)
XX
DE Full length human VEGF2 coding sequence.
XX
KW VEGF2; vascular endothelial growth factor 2; angiogenesis; bone damage;
KW endothelial cell proliferation; tissue damage; therapy; ds.
XX
XX Homo sapiens.
XX
XX US932540-A.
XX
XX 03-AUG-1999.
XX
XX 24-DEC-1997; 97US-00999811.
XX
XX 08-MAR-1994; 94US-00207550.
XX 06-JUN-1995; 95US-00465968.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Cao L, Rosen CA, Hu J;
XX WFI; 1993-443606/37.
XX P-PSDB; AAY22320.
XX Vascular endothelial growth factor 2 for wound healing and vascular repair.
XX Disclosure; Fig 1; 49pp; English.
XX This sequence encodes the vascular endothelial growth factor 2 (VEGF2),
XX of the invention. The isolated polypeptide is useful for stimulating
XX angiogenesis, by promoting the proliferation of endothelial cells, for
XX the treatment of a wound, or for the treatment of tissue or bone damage
XX
XX Sequence 1674 BP; 501 A; 383 C; 376 G; 414 T; 0 U; 0 Other;
Query Match 100.0%; Score 1674; DB 2; Length 1674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GTCTTCCACCATGCACTCGCTGGGCTTCTTCTGTGSCGTGTTCTCTGTCTGCGCGTG 60
Db 1 GTCTTCCACCATGCACTCGCTGGGCTTCTTCTGTGSCGTGTTCTCTGTCTGCGCGTG 60
QY 61 CGTGTCTCCGGGTCTCTCGAGAGCGCCCGCGCGCGCGCCCTTCAGTCCGAGCTCG 120
Db 61 CGTGTCTCCGGGTCTCTCGAGAGCGCCCGCGCGCGCGCCCTTCAGTCCGAGCTCG 120
QY 121 ACCTCTCGAGCGGAGCGCGAGCGCGCGAGCGCCACCGCTTATGCAAGCAAGATCTGG 180
Db 121 ACCTCTCGAGCGGAGCGCGAGCGCGCGAGCGCCACCGCTTATGCAAGCAAGATCTGG 180
QY 181 AGAGCAGTGTACGGTCTGTGTCCAGTGTAGTGAATCACTCATGACTGTACTCTACCCAGAT 240
Db 181 AGAGCAGTGTACGGTCTGTGTCCAGTGTAGTGAATCACTCATGACTGTACTCTACCCAGAT 240
QY 241 ATTGGAAAATGTACAAAGTGTACCTAAGGAAGAGGCTGGCAACATACAGAGAACAGG 300
Db 241 ATTGGAAAATGTACAAAGTGTACCTAAGGAAGAGGCTGGCAACATACAGAGAACAGG 300
QY 301 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGAGAGACATTTATATACAG 360
Db 301 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGAGAGACATTTATATACAG 360
QY 361 AGATCTTGAAGATATTGATTAATGAGTGGAGAAAGACTCAATGCTATGCCACGAGGCTGT 420
Db 361 AGATCTTGAAGATATTGATTAATGAGTGGAGAAAGACTCAATGCTATGCCACGAGGCTGT 420
QY 421 GTATAGATGTGGGAGGAGTGTGGAGTGCAGCAAAACACCTCTTTAAACCTCCATGTG 480
Db 421 GTATAGATGTGGGAGGAGTGTGGAGTGCAGCAAAACACCTCTTTAAACCTCCATGTG 480
QY 481 TGTCCTGTACAGATGTGGGCTGTGCTGCAATAGTGGAGGCTGCAGTGCATGACACCA 540
Db 481 TGTCCTGTACAGATGTGGGCTGTGCTGCAATAGTGGAGGCTGCAGTGCATGACACCA 540
QY 541 GCACGAGTACCTCAGCAAGACGTATTGAAATTTACAGTGCCTCTCTCAAGGCCCA 600
Db 541 GCACGAGTACCTCAGCAAGACGTATTGAAATTTACAGTGCCTCTCTCAAGGCCCA 600
QY 601 AACAGTAAACATCAGTTTGGCAATCAGACTTCTGCGGATGCTCTTAACCTGGATG 660
Db 601 AACAGTAAACATCAGTTTGGCAATCAGACTTCTGCGGATGCTCTTAACCTGGATG 660
QY 661 TTTACAGCAAGTTTCACTTCCATATTAGACGTTTCCCTGCCAGCAACACTACCACAGTGC 720
Db 661 TTTACAGCAAGTTTCACTTCCATATTAGACGTTTCCCTGCCAGCAACACTACCACAGTGC 720
QY 721 AGCAGCGAAACAAGACTGCCCAACCAATTACATGTGGAATAATCACTGAGATGCC 780
Db 721 AGCAGCGAAACAAGACTGCCCAACCAATTACATGTGGAATAATCACTGAGATGCC 780
QY 781 TGGCTCAGGAGATTTTATGTTTTCCTCGGATGCTGGAGATGACTCAACAGATGGATTCC 840
Db 781 TGGCTCAGGAGATTTTATGTTTTCCTCGGATGCTGGAGATGACTCAACAGATGGATTCC 840
QY 841 ATGACATCTGTGGACCAACAGAGAGCTGGATGAAGAGACCTGTCAAGTGTGTCTGAGAG 900
Db 841 ATGACATCTGTGGACCAACAGAGAGCTGGATGAAGAGACCTGTCAAGTGTGTCTGAGAG 900
QY 901 CGGGCTTCGGCTTCGAGCTGTGGACCCCAAGAGACTAGACAGAACTCATGCCAGT 960
Db 901 CGGGCTTCGGCTTCGAGCTGTGGACCCCAAGAGACTAGACAGAACTCATGCCAGT 960
QY 961 GTGTCTGTAAAAAACAACCTCTTCCAGCCCAATGTGGGCGCAACCGAGATTTGATGAAA 1020
Db 961 GTGTCTGTAAAAAACAACCTCTTCCAGCCCAATGTGGGCGCAACCGAGATTTGATGAAA 1020
QY 1021 ACACATGCCAGTGTGTATGTAAAAAGAACCTGCCCGCAAAATCAACCCCTAAATCCTGGAA 1080
Db 1021 ACACATGCCAGTGTGTATGTAAAAAGAACCTGCCCGCAAAATCAACCCCTAAATCCTGGAA 1080
```

```
QY 1081 AATGTGCTCTGAATGTACAGAAAGTCCACAGAAATGCTTGTATAAGGAAGAGTTCC 1140
Db 1081 AATGTGCTCTGAATGTACAGAAAGTCCACAGAAATGCTTGTATAAGGAAGAGTTCC 1140
QY 1141 ACCACCAAAACATGAGCTGTTACAGACGGCGCATGTAGCAACCCCGCAGAGGCTTGTGAGC 1200
Db 1141 ACCACCAAAACATGAGCTGTTACAGACGGCGCATGTAGCAACCCCGCAGAGGCTTGTGAGC 1200
QY 1201 CAGATTTTCAATAGTGAAGAGTGTGCTGTGCTGCTTCCCTTCATATTGCAAGACCA 1260
Db 1201 CAGATTTTCAATAGTGAAGAGTGTGCTGTGCTTCCCTTCATATTGCAAGACCA 1260
QY 1261 AAATGAGCTAAGATGTTGCTGCTTCCAGTTCATCGATTTTCTATTATGAAAACTGTGT 1320
Db 1261 AAATGAGCTAAGATGTTGCTGCTTCCAGTTCATCGATTTTCTATTATGAAAACTGTGT 1320
QY 1321 TGCCACAGTGAAGTGTCTGTGAACAGAGAGAGCCCTTGTGGTCCATGCTAAACAAGACA 1380
Db 1321 TGCCACAGTGAAGTGTCTGTGAACAGAGAGAGCCCTTGTGGTCCATGCTAAACAAGACA 1380
QY 1381 AAAGTCTGTCTTCCCTCAACCATGTGGATAAATTTACAGAAATGGAGCTGCATCTG 1440
Db 1381 AAAGTCTGTCTTCCCTCAACCATGTGGATAAATTTACAGAAATGGAGCTGCATCTG 1440
QY 1441 CAAAAGCCCTCTTGTAAAGACTGGTTTCTGCCAATGACCAACAGCAAGATTTTCCCTC 1500
Db 1441 CAAAAGCCCTCTTGTAAAGACTGGTTTCTGCCAATGACCAACAGCAAGATTTTCCCTC 1500
QY 1501 TTGTGATTTCTTTAAAGAGTACTATATAATTTATTTCCACTAAATAATTTTCTGTC 1560
Db 1501 TTGTGATTTCTTTAAAGAGTACTATATAATTTATTTCCACTAAATAATTTTCTGTC 1560
QY 1561 ATTCATTTTATAGCAACAACAAATTTGTAAACCTCACTGTGATCAATATTTTATATCAT 1620
Db 1561 ATTCATTTTATAGCAACAACAAATTTGTAAACCTCACTGTGATCAATATTTTATATCAT 1620
QY 1621 GCAAAATATGTTTAAATTAATAATGAAAATTTGTTATATAAAAAA 1674
Db 1621 GCAAAATATGTTTAAATTAATAATGAAAATTTGTTATATAAAAAA 1674

RESULT 2
AAZ10523
ID AAZ10523 standard; cDNA; 1674 BP.
XX
AC AAZ10523;
XX
DT 16-NOV-1999 (first entry)
XX
DE cDNA encoding vascular endothelial growth factor-2 (VEGF-2).
XX
KW Human vascular endothelial growth factor-2; VEGF-2;
KW vascular endothelial cell growth; endothelial cell migration;
KW angiogenesis; blood pressure; blood flow; immune system disorder;
KW immune cell; cancer; autoimmune disorder; blood protein disorder;
KW ataxia telangiectasia; common variable immunodeficiency;
KW Digorge syndrome; HIV infection; HTLV-BLV infection;
KW leukocyte adhesion deficiency syndrome; lymphopenia;
KW phagocyte bactericidal dysfunction; severe combined immunodeficiency;
KW Wiskott-Aldrich disorder; anemia; thrombocytopenia; hemoglobinuria;
KW allergy; asthma; allergic asthma; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 12..1271
FT /*tag= a
FT /product= "VEGF-2"
FT sig_peptide 12..80
FT /*tag= b
FT mat_peptide 81..1268
FT /*tag= c
XX
```


Db	1441	AAAAAGGCTCTGTAAAGACTGGTTTCTGCAATGACCAACAGCCAGATTTTCCTC	1500
Qy	1501	TTGTGATTTCTTTAAAGAAAGACTATATAATTTTCCACTAAATAATTTGTTCTGC	1560
Db	1501	TTGTGATTTCTTTAAAGAAAGACTATATAATTTTCCACTAAATAATTTGTTCTGC	1560
Qy	1561	ATTCATTTTATAGCAACAACTGGTAAACTCAGTGTGATCAATATTTTATATCAT	1620
Db	1561	ATTCATTTTATAGCAACAACTGGTAAACTCAGTGTGATCAATATTTTATATCAT	1620
Qy	1621	GCAAAATATGTTTAAATATAATGAAATTTGATTTTATATAAAAAA	1674
Db	1621	GCAAAATATGTTTAAATATAATGAAATTTGATTTTATATAAAAAA	1674
RESULT 3			
AAA52080			
ID	AAA52080 standard; cDNA; 1674 BP.		
XX	AAA52080;		
XX	AC		
XX	XX		
DT	22-DEC-2000 (first entry)		
XX	Vascular endothelial growth factor-2 (VEGF-2) coding sequence.		
DE	Vascular endothelial growth factor 2; VEGF-2; retina; angiogenesis;		
XX	treatment; injury; degeneration; photoreceptors; eye; angiod streaks;		
KW	retinitis; pigmentosa; age-related macular degeneration;		
KW	diabetic retinopathy; human; ds.		
XX	Homo sapiens.		
OS	Homo sapiens.		
XX	Key		
XX	Location/Qualifiers		
FT	12..1271		
FT	/*tag= a		
FT	/product= "Vascular endothelial growth factor-2"		
XX	WO200045835-A1.		
XX	10-AUG-2000.		
PD	07-FEB-2000; 2000WO-US0003047.		
XX	08-FEB-1999; 99US-0119179P.		
PR	12-FEB-1999; 99US-0119926P.		
PR	03-JUN-1999; 99US-012726P.		
PR	22-DEC-1999; 99US-0171505P.		
XX	(HUMA-) HUMAN GENOME SCI INC.		
PA	Rosen CA, Alderson R, Melder R, Roschke V, Ruben SM;		
PI	WPI; 2000-532862/48.		
XX	P-PSDB; AAY97144.		
DR	Treating injury or degeneration of photoreceptors comprises administering		
XX	to a subject vascular endothelial growth factor 2 (VEGF-2).		
XX	Claim 24; Fig 1a-e; 252pp; English.		
PS	Administration of vascular endothelial growth factor 2 (VEGF-2) to a		
CC	patient can be used for treating injury or degeneration of photoreceptors		
CC	associated with e.g. angiod streaks, retinitis pigmentosa, age-related		
CC	macular degeneration, diabetic retinopathy, etc. VEGF-2 promotes		
CC	angiogenesis, the formation of new blood vessels in the retina		
XX	Sequence 1674 BP; 501 A; 383 C; 376 G; 414 T; 0 U; 0 Other;		
SQ	Query Match 100.0%; Score 1674; DB 3; Length 1674;		
	Best Local Similarity 100.0%; Pred. No. 0;		
	Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		

Qy 1 GTCTCCACCATGCACTCGCTGGGCTTCTTCTCTGGGGTGTCTCTGCTCGCGCTG 60

Db 1 GTCTCCACCATGCACTCGCTGGGCTTCTTCTCTGGGGTGTCTCTGCTCGCGCTG 60

Qy 61 CGTGTCTCCCGGTCCTCGCGAGGCGCCGCGCCGCGCGCGCTTCGAGTCCGAGTCCG 120

Db 61 CGTGTCTCCCGGTCCTCGCGAGGCGCCGCGCCGCGCGCGCTTCGAGTCCGAGTCCG 120

Qy 121 ACCTCTCGAGCGCGAGCGCGAGCGCGCGCGCGCGCGCGCTTCGAGTCCGAGTCCG 180

Db 121 ACCTCTCGAGCGCGAGCGCGAGCGCGCGCGCGCGCGCGCTTCGAGTCCGAGTCCG 180

Qy 181 AGGAGCAGTTTACGCTCTGCTCCAGTGTAGATGAATCTCATGACTGTACTTACCCAGAAT 240

Db 181 AGGAGCAGTTTACGCTCTGCTCCAGTGTAGATGAATCTCATGACTGTACTTACCCAGAAT 240

Qy 241 ATTGGAATAATGTACAAAGTGTCAAGTGTAGGAAAGAGGCTGGCAACATACAGAGACAGG 300

Db 241 ATTGGAATAATGTACAAAGTGTCAAGTGTAGGAAAGAGGCTGGCAACATACAGAGACAGG 300

Qy 301 CCAACCTCAACTCAAGGACAGAGAGACTATAAATTTGCTGAGGACACATTATAATACAG 360

Db 301 CCAACCTCAACTCAAGGACAGAGAGACTATAAATTTGCTGAGGACACATTATAATACAG 360

Qy 361 AGATCTTTGAAAAGTATTGATAATGAGTGGAGAAAGACTCAATGATGCCACGGAGGTGT 420

Db 361 AGATCTTTGAAAAGTATTGATAATGAGTGGAGAAAGACTCAATGATGCCACGGAGGTGT 420

Qy 421 GTATAGATGTGGGAGAGGATTTGGAGTCCGACCAACACCTTCTTTAAACCTCCATGTG 480

Db 421 GTATAGATGTGGGAGAGGATTTGGAGTCCGACCAACACCTTCTTTAAACCTCCATGTG 480

Qy 481 TGTCCGTCTACAGATGTGGGGGTTGCTGCAATAGTGGGGCTGCAATGATGAACACCA 540

Db 481 TGTCCGTCTACAGATGTGGGGGTTGCTGCAATAGTGGGGCTGCAATGATGAACACCA 540

Qy 541 GCACGAGTACCTCAGCAAGACGTTTATTTGAAATACAGTCCCTCTCTCTCAAGGCCCA 600

Db 541 GCACGAGTACCTCAGCAAGACGTTTATTTGAAATACAGTCCCTCTCTCTCAAGGCCCA 600

Qy 601 AACCAGTAAACAAATCAGTTTTGCCAATCACACTTCTCTCCGATGATCTCTAACTGGATG 660

Db 601 AACCAGTAAACAAATCAGTTTTGCCAATCACACTTCTCTCCGATGATCTCTAACTGGATG 660

Qy 661 TTTACAGACAAAGTTCATTCATTTATAGAGCTTCCCTGCCAGCAACACTACCAAGTGC 720

Db 661 TTTACAGACAAAGTTCATTTCCATTTATAGAGCTTCCCTGCCAGCAACACTACCAAGTGC 720

Qy 721 AGGAGCGAACAAGACCTGCCACCAATACATGTGGAATATCAATCTGCAGATGCC 780

Db 721 AGGAGCGAACAAGACCTGCCACCAATACATGTGGAATATCAATCTGCAGATGCC 780

Qy 781 TGGCTCAGGAAGATTTTATGTTTTCTCGGATGCTGGAGATGACTCAACAGATGGATCC 840

Db 781 TGGCTCAGGAAGATTTTATGTTTTCTCGGATGCTGGAGATGACTCAACAGATGGATCC 840

Qy 841 ATGACATCTGGACCAACAAAGAGCTGGATGAGAGACCTGTCTGCTGTCTGCAGAG 900

Db 841 ATGACATCTGGACCAACAAAGAGCTGGATGAGAGACCTGTCTGCTGTCTGCAGAG 900

Qy 901 CGGGCTTTCCGCTGCCAGCTGTGGAACCCCAACAAAGAACTAGACAGAACTATGCCAGT 960

Db 901 CGGGCTTTCCGCTGCCAGCTGTGGAACCCCAACAAAGAACTAGACAGAACTATGCCAGT 960

Qy 961 GTGTCTGTAAAAACAACTCTTCCCGACCAATGTGGGGCCACCGAGATTTTATGAAA 1020

Db 961 GTGTCTGTAAAAACAACTCTTCCCGACCAATGTGGGGCCACCGAGATTTTATGAAA 1020

Qy 1021 ACACATGCCAGTGTGTATGTATAAAGAACTGCCCCAGAAAATCAACCCCTTAATCTGGAA 1080

Db 1021 ACACATGCCAGTGTGTATGTATAAAGAACTGCCCCAGAAAATCAACCCCTTAATCTGGAA 1080

Qy 1081 AATGTGCTGTGAAATGTATACAGAAAGTCCACAGAAATGCTTGTATAAAGGAAGATTC 1140

Db	421	GTATAGATGTGGGAGGAGTTGGAGTCGCGACAAACACCTCTCTTAAACCTCCATGTG	480
Qy	481	TGTCGGTCTACAGATGTGGGGTGTCTGCAATAGTAGGGGCTGCAGTCGATGAACACCA	540
Db	481	TGTCGGTCTACAGATGTGGGGTGTCTGCAATAGTAGGGGCTGCAGTCGATGAACACCA	540
Qy	541	GCACGAGCTACCTCAGCAAGAGCTTATTTGAAATACAGATGCTCTCTCAAGGCCCCA	600
Db	541	GCACGAGCTACCTCAGCAAGAGCTTATTTGAAATACAGATGCTCTCTCAAGGCCCCA	600
Qy	601	AACCAAGTAACTAGTTTGGCAATCACTTCCTGCCGATGCGATGCTAAACTGGATG	660
Db	601	AACCAAGTAACTAGTTTGGCAATCACTTCCTGCCGATGCGATGCTAAACTGGATG	660
Qy	661	TTTACAGACAGTTCATTTCCATTATAGACGTTCCCTGCCAGCACACTACACAGTGC	720
Db	661	TTTACAGACAGTTCATTTCCATTATAGACGTTCCCTGCCAGCACACTACACAGTGC	720
Qy	721	AGGACGCAACAGAGCTGCCCAATACATGTGGAATACATCTGCAATGTCAGATGCC	780
Db	721	AGGACGCAACAGAGCTGCCCAATACATGTGGAATACATCTGCAATGTCAGATGCC	780
Qy	781	TGGCTCAGCAAGATTTATGTTTCTCGATGCTGGAGTCACTCAACAGATGGATTCC	840
Db	781	TGGCTCAGCAAGATTTATGTTTCTCGATGCTGGAGTCACTCAACAGATGGATTCC	840
Qy	841	ATGACATCTGTGACCAACAGAGCTGGATGAAGAGACCTGTGTCAGTGTGTCTCAGAG	900
Db	841	ATGACATCTGTGACCAACAGAGCTGGATGAAGAGACCTGTGTCAGTGTGTCTCAGAG	900
Qy	901	CGGGGCTTGGCTGCGCTGTGACCCCAACAGAGCTGATGAAGAGACCTGTGTCAGT	960
Db	901	CGGGGCTTGGCTGCGCTGTGACCCCAACAGAGCTGATGAAGAGACCTGTGTCAGT	960
Qy	961	GTGCTGTAAAACAACTCTTCCAGACCAATGTGGGGCCCAACGAGATTTGATGAAA	1020
Db	961	GTGCTGTAAAACAACTCTTCCAGACCAATGTGGGGCCCAACGAGATTTGATGAAA	1020
Qy	1021	ACACATGCCAGTGTATGTAAAGAACCTGCCCCAGAAATCAACCCCTAAATCTCTGAA	1080
Db	1021	ACACATGCCAGTGTATGTAAAGAACCTGCCCCAGAAATCAACCCCTAAATCTCTGAA	1080
Qy	1081	AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAGAGTTCC	1140
Db	1081	AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAGAGTTCC	1140
Qy	1141	ACCACCAACATGCAGCTGTACAGACGGCCATGTACGACCGCCAGAGAGCGCTGTGAGC	1200
Db	1141	ACCACCAACATGCAGCTGTACAGACGGCCATGTACGACCGCCAGAGAGCGCTGTGAGC	1200
Qy	1201	CAGATTTTCATATAGTGAAGAGTGTGCTGTCCTTCATATTTGCGAAGACCAAC	1260
Db	1201	CAGATTTTCATATAGTGAAGAGTGTGCTGTCCTTCATATTTGCGAAGACCAAC	1260
Qy	1261	AAATGAGCTAAGATGTACTGTTTCCAGTTCTATCGATTTCTATTATGAAACTGTGT	1320
Db	1261	AAATGAGCTAAGATGTACTGTTTCCAGTTCTATCGATTTCTATTATGAAACTGTGT	1320
Qy	1321	TGCCACAGTGAAGTGTCTGTGAACAGAGAGACCTTTGGGTCATGTCTAACAGACA	1380
Db	1321	TGCCACAGTGAAGTGTCTGTGAACAGAGAGACCTTTGGGTCATGTCTAACAGACA	1380
Qy	1381	AAAGTGTCTTCTGCAACCTGTGATTAACCTTACAGAAATGGAGTGGAGCTCATCTG	1440
Db	1381	AAAGTGTCTTCTGCAACCTGTGATTAACCTTACAGAAATGGAGTGGAGCTCATCTG	1440
Qy	1441	CAAAAGGCTCTTGTAAAGACTGTTTCTGCCAATGACCAACAGCCAGATTTTCTCTC	1500
Db	1441	CAAAAGGCTCTTGTAAAGACTGTTTCTGCCAATGACCAACAGCCAGATTTTCTCTC	1500
Qy	1501	TTGTGATTTCTTAAAGATGACTATATATTTTCCACTTAAATATTTCTGCTG	1560
Db	1501	TTGTGATTTCTTAAAGATGACTATATATTTTCCACTTAAATATTTCTGCTG	1560

1561

ATTCAATTTTATAGCAACAACTGGTAAACTCACTGTGATCAATATTTTATATCAT

1620

1561

ATTCAATTTTATAGCAACAACTGGTAAACTCACTGTGATCAATATTTTATATCAT

1620

1621

GCAAAATATGTTTAAATATAATGAAATTTGATTTATATAAAAAA

1674

1621

GCAAAATATGTTTAAATATAATGAAATTTGATTTATATAAAAAA

1674

Qy	1561	ATTCAATTTTATAGCAACAACTGGTAAACTCACTGTGATCAATATTTTATATCAT	1620
Db	1561	ATTCAATTTTATAGCAACAACTGGTAAACTCACTGTGATCAATATTTTATATCAT	1620
Qy	1621	GCAAAATATGTTTAAATATAATGAAATTTGATTTATATAAAAAA	1674
Db	1621	GCAAAATATGTTTAAATATAATGAAATTTGATTTATATAAAAAA	1674

RESULT 5

ABK10034

ID

ABK10034 standard; cDNA; 1674 BP.

XX

ABK10034;

XX

AC

21-MAY-2002 (first entry)

XX

Human vascular endothelial growth factor 2 (VEGF-2) cDNA.

XX

Human; vascular endothelial growth factor 2; VEGF-2; gene; ss; pVGI.1;

XX

chronic limb ischaemia; myocardial ischaemia; autoimmune disorder;

XX

allergic reaction; organ rejection; inflammatory condition; arrhythmia;

XX

hyperproliferative disorder; viral infection; bacterial infection;

XX

fungal infection; parasitic infection; cardiovascular disorder; embolism;

XX

heart valve disease; aneurysm; arterial occlusive disorder; gene therapy.

XX

Homo sapiens.

XX

Key

Location/Qualifiers

XX

CDS

12..1271

/*tag= a

/product= "Human VEGF-2"

XX

WO200211769-A1.

XX

14-FEB-2002.

XX

03-AUG-2001; 2001WO-US024658.

XX

04-AUG-2000; 2000US-0223276P.

XX

(HUMA-) HUMAN GENOME SCI INC.

XX

Coleman TA;

XX

WPI; 2002-217153/27.

XX

P-PSDB; AAU76812.

XX

Isolated nucleic acid having expression vector construct with vascular

XX

endothelial growth factor-2 insert, useful for treating chronic limb

XX

ischaemia or myocardial ischaemia, autoimmune disorders and allergic

XX

conditions.

XX

Example 9; Fig 1; 241pp; English.

XX

The invention relates to an isolated nucleic acid comprising pVGI.1

XX

expression vector construct containing the vascular endothelial growth

XX

factor 2 (VEGF-2) insert. The nucleic acid is useful for producing a host

XX

cell by transducing, transforming or transfecting a host cell with the

XX

DNA and for treating a patient having chronic limb ischaemia or

XX

myocardial ischaemia, or a disease or disorder selected from autoimmune

XX

disorders (e.g. rheumatoid arthritis, dermatitis), allergic reactions or

XX

conditions (e.g. asthma), organ rejection, inflammatory conditions (e.g.

XX

Crohn's disease), hyperproliferative disorders (e.g. Gaucher's disease),

XX

diseases due to viral, bacterial, fungal or parasitic infection,

XX

cardiovascular disorders (e.g. cardiomyopathy), arrhythmia, heart valve

XX

diseases, aneurysms, arterial occlusive disorders and embolism. This

XX

sequence represents cDNA encoding human VEGF-2

XX

Sequence 1674 BP; 501 A; 383 C; 376 G; 414 T; 0 U; 0 Other;

XX

Query Match

100.0%; Score 1674; DB 6; Length 1674;

QY 961 GTGCTCTTAAACAACTCTTCCCGAGCAATGTGGGSCCAACCGAGAAATTTGATGAAA 1020
 DB 961 GTGCTCTTAAACAACTCTTCCCGAGCAATGTGGGSCCAACCGAGAAATTTGATGAAA 1020
 QY 1021 ACATATCCAGTGTGTATGTAAAGAACTGCCCCAGAAATCAACCCCTTAATCTCGAA 1080
 DB 1021 ACATATCCAGTGTGTATGTAAAGAACTGCCCCAGAAATCAACCCCTTAATCTCGAA 1080
 QY 1081 AATGCTCTGTATGTACAGAAAGTCCACAGAAATGCTTTTAAAGGAAGAGAGTCC 1140
 DB 1081 AATGCTCTGTATGTACAGAAAGTCCACAGAAATGCTTTTAAAGGAAGAGAGTCC 1140
 QY 1141 ACCACAAAATGAGCTGTACAGAGCGCCATGTACGAAACGCGCAGAGCTTTGTGAGC 1200
 DB 1141 ACCACAAAATGAGCTGTACAGAGCGCCATGTACGAAACGCGCAGAGCTTTGTGAGC 1200
 QY 1201 CAGGATTTTCATATAGTGAAGAGTGTGTGTTGTCCTTCATATTGGCAAGACAC 1260
 DB 1201 CAGGATTTTCATATAGTGAAGAGTGTGTGTTGTCCTTCATATTGGCAAGACAC 1260
 QY 1261 AATGAGCTAAGATTTGATGTTTCCAGTTTCATGATTTCTATTTATGGAATCTGTGT 1320
 DB 1261 AATGAGCTAAGATTTGATGTTTCCAGTTTCATGATTTCTATTTATGGAATCTGTGT 1320
 QY 1321 TGCCACAGTGAAGTGTGTGAAAGAGAGAGCCCTTGTGGTCCATGCTAACAAAGACA 1380
 DB 1321 TGCCACAGTGAAGTGTGTGAAAGAGAGAGCCCTTGTGGTCCATGCTAACAAAGACA 1380
 QY 1381 AATGCTCTTTCCTGACCATGTGATACTTTACAGAAATGAGTGGAGCTCATCTG 1440
 DB 1381 AATGCTCTTTCCTGACCATGTGATACTTTACAGAAATGAGTGGAGCTCATCTG 1440
 QY 1441 CAAAGGCCCTTGTAAAGAGTGGTTTTCGCAATGACCAACAGCAAGATTTTCTC 1500
 DB 1441 CAAAGGCCCTTGTAAAGAGTGGTTTTCGCAATGACCAACAGCAAGATTTTCTC 1500
 QY 1501 TTGTGATTTCTTAAAGAGTGTATATAATTTATTTCCACTAAAATATTTTCTGC 1560
 DB 1501 TTGTGATTTCTTAAAGAGTGTATATAATTTATTTCCACTAAAATATTTTCTGC 1560
 QY 1561 ATTCAATTTTATAGCAACAAATGTTGTAATCTCACTGATCAATATTTTATATCAT 1620
 DB 1561 ATTCAATTTTATAGCAACAAATGTTGTAATCTCACTGATCAATATTTTATATCAT 1620
 QY 1621 GCAAAATATGTTTAAATATAAATGAAATTTGATTTTAAATATAAATATAAATATAA 1674
 DB 1621 GCAAAATATGTTTAAATATAAATGAAATTTGATTTTAAATATAAATATAAATATAA 1674

RESULT 8

AAD49522
 ID AAD49522 standard; cDNA; 1674 BP.

AC AAD49522;
 DT 24-MAR-2003 (first entry)

XX Human vascular endothelial growth factor (VEGF-2) cDNA.
 XX Human; vascular endothelial growth factor; VEGF-2; inflammatory disease;
 KW proliferative disorder; tumour; breast; cancer; brain; prostate; colon;
 KW lymphangioma; infection; Kaposi's sarcoma; psoriasis; immunosuppressive;
 KW rheumatoid arthritis; diabetic retinopathy; gene therapy; antimicrobial;
 KW cytoskeletal; ophthalmological; autoimmune disease; gene; ss.
 OS Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 12..1271

FT /*tag= a
 FT /product= "Human VEGF-2 protein"
 FT 12..80
 FT /*tag= b

mat_peptide 81..1268
 /*tag= c
 /product= "Human mature VEGF-2 protein"
 WO200293849-A2.
 24-OCT-2002.
 12-APR-2002; 2002WO-US011404.
 13-APR-2001; 2001US-0283391P.
 07-SEP-2001; 2001US-0317600P.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Albert VR, Ruben SM, Wager RE;
 WPI; 2003-093007/08.
 P-PSDB; AAE32091.
 New vascular endothelial growth factor (VEGF)-2 antibodies, for treating, preventing or ameliorating a disease or disorder, such as inflammatory diseases, proliferative disorders, autoimmune disorders or diabetic retinopathy.
 Example 9; Page 364-366; 399pp; English.
 The invention relates to vascular endothelial growth factor (VEGF)-2 antibodies. VEGF-2 antibodies are useful for treating, preventing or ameliorating a disease or disorder, such as inflammatory diseases or disorders, proliferative disorders, tumours, tumour metastasis, breast cancer, brain cancer, prostate cancer, colon cancer, lymphangioma, an infectious disease, Kaposi's sarcoma, an autoimmune disease, rheumatoid arthritis, psoriasis, diabetic retinopathy, a disease or disorder associated with aberrant VEGF-2 (receptor) expression, or a disease or disorder associated with the lack of VEGF-2 (receptor) function. The antibody is also useful for detecting, diagnosing, prognosing, or monitoring cancers and other hyperproliferative disorders. VEGF-2 is also used in gene therapy. The present sequence is human VEGF-2 cDNA
 Sequence 1674 BP; 501 A; 383 C; 376 G; 414 T; 0 U; 0 Other;

Query Match 100.0%; Score 1674; DB 7; Length 1674;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GTCCCTTCCACATGCACTGCTGGGCTTCTCTCTGTGGCGTGTCTCTGTCGCGCTG 60
 DB 1 GTCCCTTCCACATGCACTGCTGGGCTTCTCTCTGTGGCGTGTCTCTGTCGCGCTG 60
 QY 61 CGTGTCTCCGGGTCTCTCGAGAGCGCCCGCGCGCGCGCGCTTCGAGTCCGGACTCG 120
 DB 61 CGTGTCTCCGGGTCTCTCGAGAGCGCCCGCGCGCGCGCTTCGAGTCCGGACTCG 120
 QY 121 ACCTCTCGAGACGCGAGCGCGCGCGCGCGCGCGCGCTTATGCAAGCAAGATCTGG 180
 DB 121 ACCTCTCGAGACGCGAGCGCGCGCGCGCGCGCGCTTATGCAAGCAAGATCTGG 180
 QY 181 AGGAGAGTTTACGCTGTGTCTGTCTGCTAGATGAATCATGCTGCTACTTACCAGAT 240
 DB 181 AGGAGAGTTTACGCTGTGTCTGTCTGCTAGATGAATCATGCTGCTACTTACCAGAT 240
 QY 241 ATTGGAAAAATGTACAAGTGTCTAGCTTAAGGAAAGAGGCTGGCAACATACAGAAACAGG 300
 DB 241 ATTGGAAAAATGTACAAGTGTCTAGCTTAAGGAAAGAGGCTGGCAACATACAGAAACAGG 300
 QY 301 CCAACCTCAACTCAAGGACAGAGAGACTATAAATTTGCTGCGACATTATATACAG 360
 DB 301 CCAACCTCAACTCAAGGACAGAGAGACTATAAATTTGCTGCGACATTATATACAG 360
 QY 361 AGATCTTGAATAATTTGATTAATGAGTGGAGAAAGAGACTCAATGATGCGAGAGGTGT 420
 DB 361 AGATCTTGAATAATTTGATTAATGAGTGGAGAAAGAGACTCAATGATGCGAGAGGTGT 420

Qy 421 GTATAGATGTGGGGAAGAGTTTGGAGTCGCGAACAACACCTTCTTTAAACCTCCATGTG 480
 Db 421 GTATAGATGTGGGGAAGAGTTTGGAGTCGCGAACAACACCTTCTTTAAACCTCCATGTG 480
 Qy 481 TGTCCGCTCTACAGATGCGGGGTGCTGMAATAGTGAGGGGTGAGTGATGATGAACACCA 540
 Db 481 TGTCCGCTCTACAGATGCGGGGTGCTGMAATAGTGAGGGGTGAGTGATGATGAACACCA 540
 Qy 541 GCACGAGCTACCTCAGCAAGACGTTATTTGAAATACAGTGCTCTCTCAAGGCCCA 600
 Db 541 GCACGAGCTACCTCAGCAAGACGTTATTTGAAATACAGTGCTCTCTCAAGGCCCA 600
 Qy 601 AACAGTAACAATCAGTTTTCGCAATCAACATCTTCTCGCGATGATGCTTAACTGGATG 660
 Db 601 AACAGTAACAATCAGTTTTCGCAATCAACATCTTCTCGCGATGATGCTTAACTGGATG 660
 Qy 661 TTTACACAGCAAGTTTATTCATTATTAGACCTTCCCTGCGAGCAACACTTACCACAGTGTG 720
 Db 661 TTTACACAGCAAGTTTATTCATTATTAGACCTTCCCTGCGAGCAACACTTACCACAGTGTG 720
 Qy 721 AGGAGGCAACAGACCTGCGCCCAACATTTACATGTGGAATATCATCTCCATGCTGATGCC 780
 Db 721 AGGAGGCAACAGACCTGCGCCCAACATTTACATGTGGAATATCATCTCCATGCTGATGCC 780
 Qy 781 TGGCTCAGGAAGATTTTATGTTTCTCGGATGCTGAGATGACTCAACAGATGATTTCC 840
 Db 781 TGGCTCAGGAAGATTTTATGTTTCTCGGATGCTGAGATGACTCAACAGATGATTTCC 840
 Qy 841 ATGACATCTGTGACCAACAGAGAGTGTGATGAGAGACCTGTGATGTGCTGCTGAGAG 900
 Db 841 ATGACATCTGTGACCAACAGAGAGTGTGATGAGAGACCTGTGATGTGCTGCTGAGAG 900
 Qy 901 CGGGGCTTCGGCTTCGAGCTGTGGACCCCAACAAAGTGTGACAGAACTCATGCCAGT 960
 Db 901 CGGGGCTTCGGCTTCGAGCTGTGGACCCCAACAAAGTGTGACAGAACTCATGCCAGT 960
 Qy 961 GTGTCTGTAAACAAACTCTTCCCGACCAATGTGGGCGCAACCGAGAAATTTGATGAAA 1020
 Db 961 GTGTCTGTAAACAAACTCTTCCCGACCAATGTGGGCGCAACCGAGAAATTTGATGAAA 1020
 Qy 1021 ACATGCGCAGTGTGTATGTAAAGAACCTGCCCGCAAGAAATCAACCCCTAAATCTGAAA 1080
 Db 1021 ACATGCGCAGTGTGTATGTAAAGAACCTGCCCGCAAGAAATCAACCCCTAAATCTGAAA 1080
 Qy 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTGTTAAAGGAAAGAGTTCC 1140
 Db 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTGTTAAAGGAAAGAGTTCC 1140
 Qy 1141 ACCACCAACATGACAGCTGTACAGAGCGCATGTACAGACCGCCAGAGAGGCTTGTGAGC 1200
 Db 1141 ACCACCAACATGACAGCTGTACAGAGCGCATGTACAGACCGCCAGAGAGGCTTGTGAGC 1200
 Qy 1201 CAGATTTTCATATAGTGAAGAGTGTGCTGCTGCTTCCCTTCAATTCGCAAGACCAAC 1260
 Db 1201 CAGATTTTCATATAGTGAAGAGTGTGCTGCTGCTTCCCTTCAATTCGCAAGACCAAC 1260
 Qy 1261 AATGAGCTAAGATGTGACTGTTTCCAGTTCATCGATTTCTATTATGAAAACCTGTG 1320
 Db 1261 AATGAGCTAAGATGTGACTGTTTCCAGTTCATCGATTTCTATTATGAAAACCTGTG 1320
 Qy 1321 TGCCACAGTGAAGTGTGTGAACAGAGAGACCTTGTGGGTCCTATGTAACAAGACA 1380
 Db 1321 TGCCACAGTGAAGTGTGTGAACAGAGAGACCTTGTGGGTCCTATGTAACAAGACA 1380
 Qy 1381 AAGTCTGTCTTCTCTCAACATGTGGAATAACTTTACAGAAATGGAGTGGAGCTCATCTG 1440
 Db 1381 AAGTCTGTCTTCTCTCAACATGTGGAATAACTTTACAGAAATGGAGTGGAGCTCATCTG 1440
 Qy 1441 CAAAGGCTCTTGTAAAGACTGTTTTCTGCAATGACCAACAGCCCAAGATTTTCTCTC 1500
 Db 1441 CAAAGGCTCTTGTAAAGACTGTTTTCTGCAATGACCAACAGCCCAAGATTTTCTCTC 1500

Qy 1501 TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAAAATATTTTCTGC 1560
 Db 1501 TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAAAATATTTTCTGC 1560
 Qy 1561 ATTCAATTTTATAGCAACCAATTTGGTAAACTCACTGTGATCAATATTTTATATCAT 1620
 Db 1561 ATTCAATTTTATAGCAACCAATTTGGTAAACTCACTGTGATCAATATTTTATATCAT 1620
 Qy 1621 GCAAAATATGTTTAAATATAAATGAAATTTGTTATTTATAAAAAA 1674
 Db 1621 GCAAAATATGTTTAAATATAAATGAAATTTGTTATTTATAAAAAA 1674

RESULT 9
 AAT51371
 ID AAT51371 standard; cDNA; 1674 BP.
 XX
 AC AAT51371;
 DT 23-APR-1997 (first entry)
 XX
 DE Human vascular endothelial growth factor 2 cDNA.
 XX
 KW Vascular endothelial growth factor 2; VEGF2; angiogenesis;
 KW endothelialisation; coronary bypass surgery; vascular graft surgery;
 KW agonist; antagonist; gene therapy; diagnosis; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 12..1271
 FT sig_peptide /tag= a
 FT mat_peptide /tag= b
 FT primer_bind /tag= c
 FT primer_bind complement(150..166)
 FT /tag= d
 FT /note= "5', primer binding site"
 FT primer_bind 1254..1268
 FT /tag= e
 FT primer_bind /note= "3', primer binding site"
 FT /tag= f
 FT /note= "3', primer binding site"
 XX
 FN WO9639515-A1.
 XX
 PD 12-DEC-1996.
 XX
 PP 06-JUN-1996; 96WO-US009001.
 XX
 PR 06-JUN-1995; 95US-00465968.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Hu J, Cao L;
 DR WPI; 1997-043137/04.
 DR P-PSDB; AAW11478.
 XX
 PT DNA encoding human vascular endothelial growth factor 2 - used to promote
 PT angiogenesis or endothelialisation in vascular graft surgery.
 XX
 PS Claim 1; Fig 1; 74pp; English.
 XX
 CC A cDNA clone (AAT51371) codes for human vascular endothelial growth
 CC factor 2 (VEGF2) (AAW11478), a protein structurally related to the
 CC VEGF/PDGF family that is capable of inducing angiogenesis in vivo. It was
 CC discovered in a cDNA library derived from early stage human embryo week
 CC 9. VEGF2 polynucleotides may also be obtained from adult heart or several
 CC breast cancer cell lines. VEGF2 nucleic acids can be used in the prodn.
 CC of recombinant VEGF2, as probes to detect mutations in the VEGF2, and in

CC Gene therapy to treat patients in need of VEGF2. Antisense sequences can
CC be used as VEGF2 antagonists e.g. to inhibit growth of tumours
XX
SQ Sequence 1674 BP; 502 A; 384 C; 375 G; 413 T; 0 U; 0 Other;
Query Match 99.68; Score 1667.6; DB 2; Length 1674;
Best Local Similarity 99.88; Pred. No. 0;
Matches 1670; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 GTCCTCCACCATGCTCGCTGGGCTTCTCTCTGCGGTGCTCTCTCTGCTCGCGCTG 60
Db 1 GTCCTCCACCATGCTCGCTGGGCTTCTCTCTGCGGTGCTCTCTCTGCTCGCGCTG 60
Qy 61 CGCTGCTCCGGTCTCTCGAGCGCGCGCGCGCGCGCGCGCGCTCGAGTCGGACTCG 120
Db 61 CGCTGCTCCGGTCTCTCGAGCGCGCGCGCGCGCGCGCGCGCTCGAGTCGGACTCG 120
Qy 121 ACCTCTCGGACGCGGAGCGCGACGCGGCGGAGCGCGCTTATGCAAGCAAGATCTGG 180
Db 121 ACCTCTCGGACGCGGAGCGCGACGCGGCGGAGCGCGCTTATGCAAGCAAGATCTGG 180
Qy 181 AGGAGCAGTTACGGTCTGTGTCAGTGTAGATGAACTCATGCTGTACTCTACCCAGAT 240
Db 181 AGGAGCAGTTACGGTCTGTGTCAGTGTAGATGAACTCATGCTGTACTCTACCCAGAT 240
Qy 241 ATTGGAATAATGTACAAGTGTGCTAGGAAAGAGGCTGGCAACATAACAGAGAACAGG 300
Db 241 ATTGGAATAATGTACAAGTGTGCTAGGAAAGAGGCTGGCAACATAACAGAGAACAGG 300
Qy 301 CCAACCTCAACTAAGGACAGAGAGACTATAAATTTGTCGAGCACTATAATACAG 360
Db 301 CCAACCTCAACTAAGGACAGAGAGACTATAAATTTGTCGAGCACTATAATACAG 360
Qy 361 AGATCTTGAAAGTATTGTAATGAGTGGAGAGAGCTCAATGCTGCCACGGAGGTGT 420
Db 361 AGATCTTGAAAGTATTGTAATGAGTGGAGAGAGCTCAATGCTGCCACGGAGGTGT 420
Qy 421 GTATAGATGTGGGGAAGGAGTTGGAGTTCGCGCAAAACACCTCTTTTAAACCTCCATGTG 480
Db 421 GTATAGATGTGGGGAAGGAGTTGGAGTTCGCGCAAAACACCTCTTTTAAACCTCCATGTG 480
Qy 481 TGTCCGTCTACAGATGTGGGGTGTCTGCAATAGTAGAGGGGTGCGAGTGCATGACCA 540
Db 481 TGTCCGTCTACAGATGTGGGGTGTCTGCAATAGTAGAGGGGTGCGAGTGCATGACCA 540
Qy 541 GCACGAGTCACTCAGCAAGACGTTATTGAAATACAGTGCCTCTCTCTCAAGGCCCCA 600
Db 541 GCACGAGTCACTCAGCAAGACGTTATTGAAATACAGTGCCTCTCTCTCAAGGCCCCA 600
Qy 601 AACGAGTAAATCAGATTTTCCCAATCAGCTTCTCCGCGATGCTGCTAAACTGGATG 660
Db 601 AACGAGTAAATCAGATTTTCCCAATCAGCTTCTCCGCGATGCTGCTAAACTGGATG 660
Qy 661 TTTACAGCAAGTTTCAATCCCAATTAGAGCTTCCCTGCCAGCAACACTACCAAGTGC 720
Db 661 TTTACAGCAAGTTTCAATCCCAATTAGAGCTTCCCTGCCAGCAACACTACCAAGTGC 720
Qy 721 AGGAGCGAAACAGACCTGCCCAATACATGCTGGAATAATACATCTGCAGATGCC 780
Db 721 AGGAGCGAAACAGACCTGCCCAATACATGCTGGAATAATACATCTGCAGATGCC 780
Qy 781 TGGCTCAGGAGATTTATGTTTCTCCGATGCTGAGATGACTCAACAGATGGATTC 840
Db 781 TGGCTCAGGAGATTTATGTTTCTCCGATGCTGAGATGACTCAACAGATGGATTC 840
Qy 841 ATGACATCTGTGGACCAACAGGAGCTGGATGAAGAGACCTGTGCTGTGCGAGAG 900
Db 841 ATGACATCTGTGGACCAACAGGAGCTGGATGAAGAGACCTGTGCTGTGCGAGAG 900
Qy 901 CGGGCTTCCGCTTCCAGCTGTGGACCCCAAGAACTAGACAGAACTCATGCCAGT 960
Db 901 CGGGCTTCCGCTTCCAGCTGTGGACCCCAAGAACTAGACAGAACTCATGCCAGT 960

Qy 961 GTGCTGTAAAAACAACAACTCTTCCCGAGCCAAATGTGGGGCAACCGAGAAATTTGATGAAA 1020
Db 961 GTGCTGTAAAAACAACAACTCTTCCCGAGCCAAATGTGGGGCAACCGAGAAATTTGATGAAA 1020
Qy 1021 ACACATGCCAGTGTGTATGTAAAAAGAACCTGCCCGAGAAATCAACCCCTTAATCCTGGAA 1080
Db 1021 ACACATGCCAGTGTGTATGTAAAAAGAACCTGCCCGAGAAATCAACCCCTTAATCCTGGAA 1080
Qy 1081 AATGTCCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAAAGAAAGATTC 1140
Db 1081 AATGTCCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAAAGAAAGATTC 1140
Qy 1141 ACCACCAACATGAGCTGTGTACAGCGCCATGTACGAGCGCCATGTATTTGGCAAGACAC 1200
Db 1141 ACCACCAACATGAGCTGTGTACAGCGCCATGTACGAGCGCCATGTATTTGGCAAGACAC 1200
Qy 1201 CAGGATTTTCATATAGTGAAGAGTGTGCTGTCCTTCCATATTTGGAAGAACTGTGT 1260
Db 1201 CAGGATTTTCATATAGTGAAGAGTGTGCTGTCCTTCCATATTTGGAAGAACTGTGT 1260
Qy 1261 AAATGAGCTAAGATTTGACTGTTTTCCAGTTCATCGATTTTCTATTATGGAATACTGTGT 1320
Db 1261 AAATGAGCTAAGATTTGACTGTTTTCCAGTTCATCGATTTTCTATTATGGAATACTGTGT 1320
Qy 1321 TGCCACAGTGAAGTCTGTGAACAGAGAGACCCCTTGTGGTCCATGCTAACAAAGACA 1380
Db 1321 TGCCACAGTGAAGTCTGTGTGAACAGAGAGACCCCTTGTGGTCCATGCTAACAAAGACA 1380
Qy 1381 AAAGTCTGCTTCTGTAAGAGCTGTTTCTGCAATGACCAACAGCCAGATTTTCTC 1440
Db 1381 AAAGTCTGCTTCTGTAAGAGCTGTTTCTGCAATGACCAACAGCCAGATTTTCTC 1440
Qy 1441 CAAAAGGCTCTGTTAAAGAGCTGTTTCTGCAATGACCAACAGCCAGATTTTCTC 1500
Db 1441 CAAAAGGCTCTGTTAAAGAGCTGTTTCTGCAATGACCAACAGCCAGATTTTCTC 1500
Qy 1501 TTGCTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAATAATTTGTTCTGC 1560
Db 1501 TTGCTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAATAATTTGTTCTGC 1560
Qy 1561 ATTCAATTTTATAGCAACAAATTTGGTAAACTCAGTGTGATCAATATTTTATATCAT 1620
Db 1561 ATTCAATTTTATAGCAACAAATTTGGTAAACTCAGTGTGATCAATATTTTATATCAT 1620
Qy 1621 GCAAAATATGTTTAAATATAATGAAATGTTATTTATAAAAAA 1674
Db 1621 GCAAAATATGTTTAAATATAATGAAATGTTATTTATAAAAAA 1674

RESULT 10
AAT59929
ID AAT59929 standard; cDNA; 2031 BP.
XX
AC AAT59929;
XX
DT 05-JUN-1997 (first entry)
XX
DE Human vascular endothelial growth factor-related protein VRP cDNA.
XX
KW Vascular endothelial growth factor-related protein; VEGF;
KW receptor protein tyrosine kinase; Flt4; signal transduction;
KW wound healing; vulvar; rheumatoid arthritis; Kaposi's sarcoma;
KW therapy; diagnosis; angiogenesis; ds.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH 372..1631
FT CDS /*tag= a
FT /transl_except= pos:711..713_aa:Thr
FT sig_peptide 372..431 /*tag= b
FT mat_peptide 432..1628

FT		/ *tag= c	
XX	PN	WO9709427-A1.	
XX	PD	13-MAR-1997.	
XX	PF	30-AUG-1996; 96WO-US014075.	
XX	PR	08-SEP-1995; 95US-0003491P.	
XX	PA	(GETH) GENENTECH INC.	
XX	PI	Lee J, Wood W;	
XX	DR	WPI; 1997-192902/17.	
XX	DR	P-FSDB; AAW13933.	
XX	PT	Human protein similar to vascular endothelial growth factor - used to treat e.g. wounds, tumours, rheumatoid arthritis, Kaposi's sarcoma etc.	
XX	PS	Example 3; Fig 1A-D; 68pp; English.	
CC	CC	A cdna clone (AAT59929), designated VH1.4, codes for human vascular endothelial growth factor (VEGF)-related protein (VRP) (AAW13933), a protein that binds to, and stimulates the phosphorylation of, the receptor tyrosine kinase Flt4. To obtain VH1.4, a chimera confg. the extracellular domain of the Flt4 receptor and the Fc region of the IgG heavy chain was expressed in 293 cells and used to screen for membrane-linked VRP by fluorescent-activated cell sorting. A positive cell line, Glioma G6i, was identified and a cDNA library from it was screened with probes (see also AAT59931-32) based on an EST sequence (AAT59930)	
CC	CC	sequence similar to VEGF to obtain clone VH1.4. The clone can be used to produce recombinant VRP, and diagnostically for tissue-specific typing	
XX	SQ	Sequence 2031 BP; 541 A; 532 C; 499 G; 459 T; 0 U; 0 Other;	
		Query Match 99.4%; Score 1664.6; DB 2; Length 2031;	
		Best Local Similarity 99.8%; Pred. No. 0;	
		Matches 1667; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
Qy	1	GTCCTTCCACCATGCAGCTCGCTGGGCTTCTCTGTGGCGTGTTCTCTGCTCGCCGCTG 60	
Db	361	GTCCTTCCACCATGCAGCTTGTGGGCTTCTCTGTGGCGTGTTCTCTGCTCGCCGCTG 420	
Qy	61	CGTGTCTCCGGGTCTCTCGCAGGCCGCGCGCGCGCGCGCTTCGAGTCCGACTCG 120	
Db	421	CGTGTCTCCGGGTCTCTCGCAGGCCGCGCGCGCGCGCGCTTCGAGTCCGACTCG 480	
Qy	121	ACCTCTCGGACGGGAGCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180	
Db	481	ACCTCTCGGACGGGAGCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540	
Qy	181	AGGAGCAGTTACGGTCTGTGTCCAGTGTAGTAGTAACCTCATGACTGTACTCTACCCAGAAT 240	
Db	541	AGGAGCAGTTACGGTCTGTGTCCAGTGTAGTAGTAACCTCATGACTGTACTCTACCCAGAAT 600	
Qy	241	ATTGGAATAATGACAGTGTACGTTAGGTAAGGAAGGCGGTGCGACATTAACAAGAACAGG 300	
Db	601	ATTGGAATAATGACAGTGTACGTTAGGTAAGGAAGGCGGTGCGACATTAACAAGAACAGG 660	
Qy	301	CCAACCTCAACTCAAGGACAGAGAGACTATAAAAATTTGTGCGAGCACATTATAATACAG 360	
Db	661	CCAACCTCAACTCAAGGACAGAGAGACTATAAAAATTTGTGCGAGCACATTATAATACAG 720	
Qy	361	AGATCTTGAAGATTTGATATGATGAGGAGAGACTCAATGATGCCACGGGAGGTG 420	
Db	721	AGATCTTGAAGATTTGATATGATGAGGAGAGACTCAATGATGCCACGGGAGGTG 780	
Qy	421	GTATAGATGTGGGAGGAGTTTGGAGTCGGGACAAACACCTTCTTTAAACCTCCATGTG 480	
Db	781	GTATAGATGTGGGAGGAGTTTGGAGTCGGGACAAACACCTTCTTTAAACCTCCATGTG 840	
Qy	481	TGTCCTCTACAGATGTGGGGGTTGCTGCAATATGATGAGGGGTCGATGTCATGAACACCA 540	
Db	841	TGTCGCTCTACAGATGTGGGGTGTCTGCAATATGATGAGGGGCTCGATGTCATGTAACCA 900	
Qy	541	GCACGAGCTACCTCAGCAAGACGTTATTGAAATTAACAGTGCCCTCTCTCAAGGCCCA 600	
Db	901	GCACGAGCTACCTCAGCAAGACGTTATTGAAATTAACAGTGCCCTCTCTCAAGGCCCA 960	
Qy	601	AAACAGTAACATCAGTTTTGCCAATCACACTTCTCCGATGCATCTCTAAACTGGATG 660	
Db	961	AAACAGTAACATCAGTTTTGCCAATCACACTTCTCCGATGCATCTCTAAACTGGATG 1020	
Qy	661	TTTACAGCAAGTTCATTTCCATTTAGACGTTCCCTGCCAGCAACACTACCACAGTGC 720	
Db	1021	TTTACAGCAAGTTCATTTCCATTTAGACGTTCCCTGCCAGCAACACTACCACAGTGC 1080	
Qy	721	AGGCAGCGAACAGACCTGCCCAACCAATTACATGTGGAATATCAATCTGCAGATGCC 780	
Db	1081	AGGCAGCGAACAGACCTGCCCAACCAATTACATGTGGAATATCAATCTGCAGATGCC 1140	
Qy	781	TGCTCAGCAAGATTTTATGTTTTCTCTCGATGCTGGAGATGACTCAAACAGATGGATTC 840	
Db	1141	TGCTCAGCAAGATTTTATGTTTTCTCTCGATGCTGGAGATGACTCAAACAGATGGATTC 1200	
Qy	841	ATGACATCTGTGGACCAACAGAGAGCTGCATGAGAGACTCTGAGTGTGTCGAGAG 900	
Db	1201	ATGACATCTGTGGACCAACAGAGAGCTGCATGAGAGACTCTGAGTGTGTCGAGAG 1260	
Qy	901	CGGGCTTTCGGCTCGCCAGCTGTGGACCCCAACAAAGAACTAGACAGAAACTCATGCCAGT 960	
Db	1261	CGGGCTTTCGGCTCGCCAGCTGTGGACCCCAACAAAGAACTAGACAGAAACTCATGCCAGT 1320	
Qy	961	GTGCTGTFAAAAAAACTCTTCCCAGCCAAATGTGGGGCCCAACAGAGAAATTTGATGAAA 1020	
Db	1321	GTGCTGTFAAAAAAACTCTTCCCAGCCAAATGTGGGGCCCAACAGAGAAATTTGATGAAA 1380	
Qy	1021	ACATGCGCAGTGTGTATGTAAAAAGAACCTGCCGCCAGAAATCAACCCCTAAATCCCTGAAA 1080	
Db	1381	ACATGCGCAGTGTGTATGTAAAAAGAACCTGCCGCCAGAAATCAACCCCTAAATCCCTGAAA 1440	
Qy	1081	AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTGTTAAAGGAAGAAGTTCC 1140	
Db	1441	AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTGTTAAAGGAAGAAGTTCC 1500	
Qy	1141	ACCACCAAAACATGCAGCTGTTACAGACGGCCATGTACGAACCGCCAGAGAGGCTTGTGAGC 1200	
Db	1501	ACCACCAAAACATGCAGCTGTTACAGACGGCCATGTACGAACCGCCAGAGAGGCTTGTGAGC 1560	
Qy	1201	CAGATTTTCAATATAGTGAAGAGTGTGTGTGTCTCCCTTCATATTTGGCAAGACCAAC 1260	
Db	1561	CAGATTTTCAATATAGTGAAGAGTGTGTGTGTCTCCCTTCATATTTGAAAGAACCAAC 1620	
Qy	1261	AAATGAGCTAAAGATTGTACTGTTTCCAGTTTCATTCGATTTTCTATTATGAAAACTGTGT 1320	
Db	1621	AAATGAGCTAAAGATTGTACTGTTTCCAGTTTCATTCGATTTTCTATTATGAAAACTGTGT 1680	
Qy	1321	TGCCACAGTGAACCTGTCTGTGAACAGAGAGACCCCTTGTGGGTCCATGCTATCAAAAGACA 1380	
Db	1681	TGCCACAGTGAACCTGTCTGTGAACAGAGAGACCCCTTGTGGGTCCATGCTATCAAAAGACA 1740	
Qy	1381	AAAGTCTGTCTTTCCTGAACCATGTGGATACTTTACAGAAATGAGCTGAGCTCATCTG 1440	
Db	1741	AAAGTCTGTCTTTCCTGAACCATGTGGATACTTTACAGAAATGAGCTGAGCTCATCTG 1800	
Qy	1441	CAAAAGGCTCTTGTAAAGACTGGTTTTCTGCCAATGACCAACAGCCAAAGATTTTCCCTC 1500	
Db	1801	CAAAAGGCTCTTGTAAAGACTGGTTTTCTGCCAATGACCAACAGCCAAAGATTTTCCCTC 1860	
Qy	1501	TTTGATGATTTTAAAGATGACTATATAAATTTATTTTCC	

Db 1921 ATTCAATTTTATAGCAACACAAATTCGTTAAACTCACTGTGATCAATATTTTATATCAT 1980

Qy 1621 GCATAATATGTTAAATAAATAAGAAATGTATTATTAATAAAAAAAAAAAAA 1671

Db 1981 GCAAAATATGTTTAAATAAATAAGAAATGTATTATTAATAAAAAAAAAAAAA 2031

RESULT 11

AAT84276

ID AAT84276 standard; cDNA; 1997 BP.

XX AAT84276;

XX

DT 10-NOV-1997 (first entry)

XX Human Flt4 receptor tyrosine kinase ligand VEGF-C cDNA.

DE

XX VEGF-C; Flt4; receptor tyrosine kinase; VEGFR-3; human;

KW vascular endothelial growth factor receptor-3; ligand; angiogenesis;

KW wound healing; lymph vessel; lymphangioma; cancer; metastasis; therapy;

KW diagnosis; plasmid pFLT-4-L; ss.

XX Homo sapiens.

XX

XX Key Location/Qualifiers

FT CDS 352..1611

FT /*tag= a

FT sig_peptide 352..657

FT /*tag= b

FT mat_peptide 658..1608

FT /*tag= c

XX

PN WO9705250-A2.

XX

PD 13-FEB-1997.

XX

XX 01-AUG-1996; 96WO-FI000427.

XX

PR 01-AUG-1995; 95US-00510133.

PR 12-JAN-1996; 96US-00585895.

PR 14-FEB-1996; 96US-00601132.

PR 28-JUN-1996; 96US-00671573.

XX

XX (UYHE-) UNIV HELSINKI LICENSING LTD OY.

XX

PI Alitalo K, Joukov V;

XX

DR WPI; 1997-145688/13.

DR P-PSDB; AAW00932.

XX

XX Flt4 receptor tyrosine kinase ligand and related nucleic acid - used to

PT modulate growth of endothelial cells and for diagnosis of endothelial

PT cell diseases.

XX

XX Claim 25; Page 110-112; 183pp; English.

XX

XX This cDNA clone codes for human VEGF-C (AAW00932), a novel ligand that

CC binds specifically to human Flt4 receptor tyrosine kinase (VEGFR-3),

CC stimulating phosphorylation of the receptor, and which can be used to

CC modulate growth of endothelial cells and for diagnosis of endothelial

CC cell diseases. The clone was isolated from a PC-3 prostatic

CC adenocarcinoma cell (ATCC CRL 1435) cDNA library by PCR amplification and

CC has been cloned into pCDNA1 to form plasmid pFLT4-L (ATCC 97231). The

CC human VEGF-C gene is located on chromosome 4q34 and contains 6 introns

CC (see A184303-14). The insert from the human clone was used to isolate

CC mouse (see AAR84277) and quail (see AAR84300) VEGF-C sequences. It can

CC also be used to produce recombinant VEGF-C polypeptides in host cells and

CC to provide probes for screening endothelial cell diseases. Antisense

CC molecules that block the receptor can be used to control endothelial cell

CC proliferation, e.g. lymphangioma or metastatic cancer. The promoter

XX sequence (AAR84278) is also claimed

XX

XX Sequence 1997 BP; 526 A; 527 C; 490 G; 454 T; 0 U; 0 Other;

SQ

Query Match 98.7%; Score 1651.8; DB 2; Length 1997;

Best local similarity 99.9%; Pred. No. 0;

Matches 1653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GTCTTCCACCATGCACTGCTGGGCTTCTTCTGTGGCGGTGTTCTCTGCTGGCGCTG 60

Db 341 GTCTTCCACCATGCACTGCTGGGCTTCTTCTGTGGCGGTGTTCTCTGCTGGCGCTG 400

Qy 61 CGTGTCTCCGGTCTCTCGAGGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120

Db 401 CGTGTCTCCGGTCTCTCGAGGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 460

Qy 121 ACCTCTCGGACCGGAGCCGAGCGCGGCGAGGCGCGCGCGCGCGCGCGCGCGCGCG 180

Db 461 ACCTCTCGGACCGGAGCCGAGCGCGGCGAGGCGCGCGCGCGCGCGCGCGCGCGCG 520

Qy 181 AGGAGCAGTTACGCTGTGTCTCAGTGTAGATGAACCTCATGACTGTACTCTACCCAGA 240

Db 521 AGGAGCAGTTACGCTGTGTCTCAGTGTAGATGAACCTCATGACTGTACTCTACCCAGA 580

Qy 241 ATTGGAAATGTACAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCA 300

Db 581 ATTGGAAATGTACAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCAAGTGTCA 640

Qy 301 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGCAGCACATTTATAATAC 360

Db 641 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGCAGCACATTTATAATAC 700

Qy 361 AGATCTTGAAGATTTGATTAATGATGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420

Db 701 AGATCTTGAAGATTTGATTAATGATGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 760

Qy 421 GTATAGATGTGGGAGGAGTTTGGAGTGCAGACAAACACCTTCTTTAAACCTCCATGTG 480

Db 761 GTATAGATGTGGGAGGAGTTTGGAGTGCAGACAAACACCTTCTTTAAACCTCCATGTG 820

Qy 481 TGTCCGTCTACAGATGTGGGAGTTTGTGCAATAGTGTGAGGGGTGAGTGCATGAACACCA 540

Db 821 TGTCCGTCTACAGATGTGGGAGTTTGTGCAATAGTGTGAGGGGTGAGTGCATGAACACCA 580

Qy 541 GCACGAGCTACCTCAGCAGAGGTTTATTGAAATTTACAGTGCCTCTCTCTGAGGCCCA 600

Db 881 GCACGAGCTACCTCAGCAGAGGTTTATTGAAATTTACAGTGCCTCTCTCTGAGGCCCA 940

Qy 601 AACCGATTAACAATCAGTTTGTCCCAATCACCTTCTCTGCGGATGCATGTCTAACTGGATG 660

Db 941 AACCGATTAACAATCAGTTTGTCCCAATCACCTTCTCTGCGGATGCATGTCTAACTGGATG 1000

Qy 661 TTACAGACAGTTTCAATTCATTTAGAGTGTCCCTGCGAGCAACACTACACAGTGTG 720

Db 1001 TTACAGACAGTTTCAATTCATTTAGAGTGTCCCTGCGAGCAACACTACACAGTGTG 1060

Qy 721 AGGACGCAACAGACCTGCCCCACCAATTACATGTGGAATAATCACATCTGCAGATGCC 780

Db 1061 AGGACGCAACAGACCTGCCCCACCAATTACATGTGGAATAATCACATCTGCAGATGCC 1120

Qy 781 TGGCTCAGGAAGATTTATGTTTCTCGGATGTGGAGATGACTCAACAGATGATTC 840

Db 1121 TGGCTCAGGAAGATTTATGTTTCTCGGATGTGGAGATGACTCAACAGATGATTC 1180

Qy 841 ATGACATCTGTGGACCAACAGAGGAGCTGGATGAGAGACCTGTGCTGCTGCAGAG 900

Db 1181 ATGACATCTGTGGACCAACAGAGGAGCTGGATGAGAGACCTGTGCTGCTGCAGAG 1240

Qy 901 CGGGGCTTCGGCTCGCCAGCTGTGTGACCCCAACAAAGAACTAGACAGAAACTCATGCCAGT 960

Db 1241 CGGGGCTTCGGCTCGCCAGCTGTGTGACCCCAACAAAGAACTAGACAGAAACTCATGCCAGT 1300

Qy 961 GTGTCTGTAATAAACAACACTTTCGCCAGCAATGTGGGGCCAAACCGAGAAATTTGATGAA 1020

Db 1301 GTGTCTGTAATAAACAACACTTTCGCCAGCAATGTGGGGCCAAACCGAGAAATTTGATGAA 1360

1021 ACACATGCCAGTGTATGTAAAGAACCTGCCCGAGAAATCAACCCCTAAATCCTGGAA 1080
 Db ACACATGCCAGTGTATGTAAAGAACCTGCCCGAGAAATCAACCCCTAAATCCTGGAA 1420
 Qy AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGAAAGAGATTCC 1140
 Db AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGAAAGAGATTCC 1480
 Qy ACCACCAACATGAGCTGTACAGAGCGGCATGTACAGCCCGCAGAGCCTGTGAGC 1200
 Db ACCACCAACATGAGCTGTGTACAGAGCGGCATGTACAGCCCGCAGAGCCTGTGAGC 1540
 Qy CAGGATTTTCATATAGTGAAGAGTGTGCTGTGTCCTTTCATATTTGGCAAGACCAC 1260
 Db CAGGATTTTCATATAGTGAAGAGTGTGCTGTGTCCTTTCATATTTGGCAAGACCAC 1600
 Qy AATGAGCTAAGATTGCTTTTCCAGTTTCATCGATTTTCTATTATGGAATCTGTGT 1320
 Db AATGAGCTAAGATTGCTTTTCCAGTTTCATCGATTTTCTATTATGGAATCTGTGT 1660
 Qy TGCCACAGTAACTGCTGTGAACAGAGAGCCCTTGTGGTCCATGCTTAAAGAGCA 1380
 Db TGCCACAGTAACTGCTGTGAACAGAGAGCCCTTGTGGTCCATGCTTAAAGAGCA 1720
 Qy AAGTCTGTCTTCTCTGAACTGTGGATAACTTTACAGAAATGGAGCTGAGCTCATCTG 1440
 Db AAGTCTGTCTTCTCTGAACTGTGGATAACTTTACAGAAATGGAGCTGAGCTCATCTG 1780
 Qy CAAAAGCCCTCTGTAAAGACTGTTTCTGCCAATGACCAACAGCCCAAGATTTCCTC 1500
 Db CAAAAGCCCTCTGTAAAGACTGTTTCTGCCAATGACCAACAGCCCAAGATTTCCTC 1840
 Qy TTGTGATTTCTTTAAAGATGACTATATAATTTATTTCCACTTAAATAATTTCTGC 1560
 Db TTGTGATTTCTTTAAAGATGACTATATAATTTATTTCCACTTAAATAATTTCTGC 1900
 Qy ATTCTATTTTATAGCAACAAATTTGGTAAACTCACTGTGATCAATATTTTATATCAT 1620
 Db ATTCTATTTTATAGCAACAAATTTGGTAAACTCACTGTGATCAATATTTTATATCAT 1960
 Qy GCAAAATATGTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1655
 Db GCAAAATATGTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1995

RESULT 12
 AAV52576
 ID AAV52576 standard; DNA; 1997 BP.
 AC AAV52576;
 XX 14-DEC-1998 (first entry)
 DE Human vascular endothelial growth factor C gene.
 KW Flt4; vascular endothelial growth factor C; vascular endothelial cell;
 KW lymphatic endothelial cell; myelopoiesis; angiogenesis; inflammation;
 KW lymphangiogenesis; oedema; elephantiasis; Milroy's disease; ss.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FH 352..1611
 FT /*tag= a
 FT /product= "VEGF-C protein"
 XX
 PN W09833917-A1.
 XX
 PD 06-AUG-1998.
 XX
 PF 02-FEB-1998; 98MO-US001973.
 XX
 PR 05-FEB-1997; 97US-00795430.

XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 PA (UYHE-) UNIV HELSINKI LICENSING LTD.
 XX
 PI Alitalo K, Joukov V;
 XX
 DR WPI; 1998-437470/37.
 DR P-PSDB; AAW75740.
 XX
 PT New isolated vascular endothelial growth factor polypeptide(s) - used to
 PT develop products for treating, e.g. cancers, inflammation, oedema,
 PT granulocytopenia or for wound healing or tissue transplantation.
 XX
 PS Example 10; Page 112-115; 177pp; English.
 XX
 CC The vascular endothelial growth factor C (VEGF-C) polypeptides have
 CC activities affecting growth and migration of vascular endothelial cells,
 CC promoting growth of lymphatic endothelial cells and lymphatic vessels,
 CC increasing vascular permeability, and affecting myelopoiesis. The
 CC products can be used for stimulating angiogenesis, for inhibiting
 CC angiogenesis, for stimulating lymphangiogenesis, treatment or prevention
 CC of inflammation, oedema, elephantiasis, or Milroy's disease. They can
 CC also be used to modulate myelopoiesis, e.g. treating granulocytopenia.
 CC They can also be used for modulating the growth of endothelial cells.
 CC They can also be used to stimulate lymphocyte production and maturation,
 CC and to promote or inhibit trafficking of leucocytes between tissues and
 CC lymphatic vessels or to affect migration in and out of the thymus
 XX
 SQ Sequence 1997 BP; 526 A; 527 C; 490 G; 454 T; 0 U; 0 Other;
 Query Match 98.7%; Score 1651.8; DB 2; Length 1997;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GTCCTTCCACCATGCTGCTGGCTTCTTCTGTGGGTGTCTCTGCTCCCGCTG 60
 Db 341 GTCCTTCCACCATGCTGCTGGCTTCTTCTGTGGGTGTCTCTGCTCCCGCTG 400
 QY 61 CGCTGCTCCCGGTCTCGGAGCGCCGCGCGCGCGCGCGCGCTTCGAGTCGGACTCG 120
 Db 401 CGCTGCTCCCGGTCTCGGAGCGCCGCGCGCGCGCGCTTCGAGTCGGACTCG 460
 QY 121 ACCTCTCGAGCGCGAGCGCGCGCGCGCGCGCGCTTATGCAAGCAAGATCTGG 180
 Db 461 ACCTCTCGAGCGCGAGCGCGCGCGCGCGCGCTTATGCAAGCAAGATCTGG 520
 QY 181 AGGAGCAGTTACGTTCTGTGTCAGTGTAGTGAATCATGACTGTACTTATCCAGAT 240
 Db 521 AGGAGCAGTTACGTTCTGTGTCAGTGTAGTGAATCATGACTGTACTTATCCAGAT 580
 QY 241 ATTGGAATAATGTACAAGTGTCAAGTAAAGAGAGGCTGGCAACATAACAGAGAACAG 300
 Db 581 ATTGGAATAATGTACAAGTGTCAAGTAAAGAGAGGCTGGCAACATAACAGAGAACAG 640
 QY 301 CCAACTCTAACTCAAGGACAGAGAGACTATAAAATTTGCTGAGCAGACATTAATATACAG 360
 Db 641 CCAACTCTAACTCAAGGACAGAGAGACTATAAAATTTGCTGAGCAGACATTAATATACAG 700
 QY 361 AGATCTTGAAGTATTTGATTAATGAGTGGAGAGAGACTCAATGCATGCCACGGAGGTGT 420
 Db 701 AGATCTTGAAGTATTTGATTAATGAGTGGAGAGAGACTCAATGCATGCCACGGAGGTGT 760
 QY 421 GTATAGATGTGGGAAAGGAGTTTGGAGTGGCAACAAACCTTCTTTTAAACCTCCATGTG 480
 Db 761 GTATAGATGTGGGAAAGGAGTTTGGAGTGGCAACAAACCTTCTTTTAAACCTCCATGTG 820
 QY 481 TGTCCGTCTACAGATGTGGGAGTTTGTGCAATAGTGGGGGTGCAAGTGCATGAACACCA 540
 Db 821 TGTCCGTCTACAGATGTGGGAGTTTGTGCAATAGTGGGGGTGCAAGTGCATGAACACCA 880
 QY 541 GCACGAGCTACCTCAGCAAGACGTTATTTTGAATTAACAGTGCCTCTCTCTCAAGGCCCA 600
 Db 881 GCACGAGCTACCTCAGCAAGACGTTATTTTGAATTAACAGTGCCTCTCTCTCAAGGCCCA 940

Qy 601 AACAGTAACAATCAGTTTTCCTCCAAATCACATTTCTCTCCGATGCAITGCTTAAACTGGATG 660
Db 941 AACAGTAACAATCAGTTTTCCTCCAAATCACATTTCTCTCCGATGCAITGCTTAAACTGGATG 1000
Qy 661 TTTACACAGAGTTTCAATTCATTTAGAGTTTCCCTGCGCAGCAACACTACCACTGTC 720
Db 1001 TTTACACAGAGTTTCAATTCATTTAGAGTTTCCCTGCGCAGCAACACTACCACTGTC 1060
Qy 721 AGCAGCGAACAAGACCTGCCCAACCAATTTACATGTGGAAATACATCTGCAGATGCC 780
Db 1061 AGCAGCGAACAAGACCTGCCCAACCAATTTACATGTGGAAATACATCTGCAGATGCC 1120
Qy 781 TGGCTCAGGAAGATTTTATGTTTCTCCGATGCTGAGATGACTCAACAGATGATTC 840
Db 1121 TGGCTCAGGAAGATTTTATGTTTCTCCGATGCTGAGATGACTCAACAGATGATTC 1180
Qy 841 ATGACATCTGTGGACCAAAACAAGAGCTGGATGAAGAGACCTGCTAGTGTGCTGCAGAG 900
Db 1181 ATGACATCTGTGGACCAAAACAAGAGCTGGATGAAGAGACCTGCTAGTGTGCTGCAGAG 1240
Qy 901 CGGGCTTCGGCTCGCCAGCTGTGGACCCCAAGAACTAGACAGAACTCATGCCAGT 960
Db 1241 CGGGCTTCGGCTCGCCAGCTGTGGACCCCAAGAACTAGACAGAACTCATGCCAGT 1300
Qy 961 GTGCTGTAAACAAACTCTTCCCAAGCAATGTGGGCGCAACCGAGAAATTTGATGAA 1020
Db 1301 GTGCTGTAAACAAACTCTTCCCAAGCAATGTGGGCGCAACCGAGAAATTTGATGAA 1360
Qy 1021 ACATGCCAGTGTGATGTAAAGAACTGCCCCAGAAATCAACCCCTAAATCTCTGGAA 1080
Db 1361 ACATGCCAGTGTGATGTAAAGAACTGCCCCAGAAATCAACCCCTAAATCTCTGGAA 1420
Qy 1081 AATGTGCTGTGAATGTACAGAAAGTCCAGAAATGCTGTGTTAAAGAAAGTGTCC 1140
Db 1421 AATGTGCTGTGAATGTACAGAAAGTCCAGAAATGCTGTGTTAAAGAAAGTGTCC 1480
Qy 1141 ACCACCAACATGCTGTTACAGACGGCGCATGTACAGACCCCGCAAGAGGTTGTGAGC 1200
Db 1481 ACCACCAACATGCTGTTACAGACGGCGCATGTACAGACCCCGCAAGAGGTTGTGAGC 1540
Qy 1201 CAGATTTTATATAGTGAAGAGTGTGCTGTGCTCCCTTCATATTTGGCAAGACAC 1260
Db 1541 CAGATTTTATATAGTGAAGAGTGTGCTGTGCTCCCTTCATATTTGGCAAGACAC 1600
Qy 1261 AAATGAGCTAAGATTTGATGTTTCCAGTTCATCGATTTTCTATTATGCAAACTGTGT 1320
Db 1601 AAATGAGCTAAGATTTGATGTTTCCAGTTCATCGATTTTCTATTATGCAAACTGTGT 1660
Qy 1321 TGCCACAGTGAAGTGTGTAACAGAGAGACCTTTGGGGTCCATGCTAACAAAGACA 1380
Db 1661 TGCCACAGTGAAGTGTGTAACAGAGAGACCTTTGGGGTCCATGCTAACAAAGACA 1720
Qy 1381 AAAGTGTGCTTCTTCCGAACATGTGATTAACATTTACAGAAATGGAGCTGATCTG 1440
Db 1721 AAAGTGTGCTTCTTCCGAACATGTGATTAACATTTACAGAAATGGAGCTGATCTG 1780
Qy 1441 CAAAAGCCCTCTTGTAAAGACTGGTTTCTGCCAATGACCAACAGCCCAAGATTTTCCTC 1500
Db 1781 CAAAAGCCCTCTTGTAAAGACTGGTTTCTGCCAATGACCAACAGCCCAAGATTTTCCTC 1840
Qy 1501 TTGTGATTTTAAAGAAATGACTATATAATTTATTTCCACTAAATATTTGTTCTGC 1560
Db 1841 TTGTGATTTTAAAGAAATGACTATATAATTTATTTCCACTAAATATTTGTTCTGC 1900
Qy 1561 ATTCAATTTTATAGCAACAATTTGGTAAATCTCACTGTGATCAATATTTTATATCAT 1620
Db 1901 ATTCAATTTTATAGCAACAATTTGGTAAATCTCACTGTGATCAATATTTTATATCAT 1960
Qy 1621 GCAAAATATGTTTAAATATAAATGAAATGTATT 1655
Db 1961 GCAAAATATGTTTAAATATAAATGAAATGTATT 1995

RESULT 13
AAD00339
ID AAD00339 standard; cDNA; 1997 BP.
XX
AC AAD00339;
XX
DT 09-AUG-2000 (first entry)
XX
DE Human vascular endothelial growth factor (VEGF)-C cDNA.
XX
KW Vascular endothelial growth factor-C; VEGF; human; re-endothelialisation;
KW Vascular endothelial growth factor receptor; VEGFR; vascular trauma;
KW blood vessel; cardiovascular surgery; anti-restenosis agent; prevention;
KW restenosis; stenosis; percutaneous transluminal coronary angioplasty; ss.
OS Homo sapiens.
XX
FH Key
FT CDS
FT
FT /tag= a
FT /product= "Human prepro-VEGF-C protein"
FT /note= "Vascular endothelial growth factor-C"
FT sig_peptide 352..444
FT /tag= b
FT /note= "Cleavage results in partially-processed VEGF-C protein"
FT misc_feature 445..660
FT /tag= c
FT /note= "Encodes amino terminal peptide, which on cleavage results in fully-processed VEGF-C protein"
FT mat_peptide 661..1032
FT /tag= d
FT /product= "Mature human VEGF-C protein"
FT /note= "Processed and active vascular endothelial growth factor-C"
FT misc_feature 1033..1608
FT /tag= e
FT /note= "Encodes carboxy terminal peptide, which on cleavage results in partially-processed VEGF-C protein"
XX
FN WO200024412-A2.
XX
XX
PD 04-MAY-2000.
XX
PF 26-OCT-1999; 99WO-US024054.
XX
PR 26-OCT-1998; 98US-0105587P.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
PA (UYHE-) UNIV HELSINKI LICENSING LTD OY.
PA (YLAH/) YLA-HERTTUALA S.
XX
PI Yla-Herttuala S, Alitalo K, Hiltunen MC, Jeltsch MW, Achen MG;
XX
XX WPI; 2000-350584/30.
DR P-PSDB; AAY0982.
XX
PT Preventing stenosis and restenosis in mammals using vascular endothelial growth factor proteins or the nucleic acids encoding them.
XX
PS Example 1; Page 49-51; 61pp; English.
XX
CC The present cDNA sequence encodes the complete human prepro-vascular endothelial growth factor (VEGF)-C. VEGF-C has the ability to stimulate re-endothelialisation of an injured blood vessel, without significant stimulation of smooth muscle cell proliferation. It can bind to and stimulate VEGFR-2 (vascular endothelial growth factor receptor) and/or VEGFR-3 phosphorylation in cells that express such receptors. An anti-restenosis agent comprising either a VEGF-C gene or protein is used in a method to reduce or prevent restenosis and stenosis of a blood vessel following vascular trauma e.g., cardiovascular surgery and percutaneous transluminal coronary angioplasty

XX	SQ	Sequence	1997 BP;	526 A;	527 C;	490 G;	454 T;	0 U;	0 Other;
		Query Match Best Local Similarity 98.7%; Score 1651.8; DB 3; Length 1997; Matches 1653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;							
QY	1	GTCCTCCACCATGCACTCGTGGGGTTCTTCTGTGGGGTGTTCTCTGCCCGCTG 60							
DB	341	GTCTCTCCACCATGCATCTGCTGGGGTTCTTCTGTGGGGTGTTCTCTGCCCGCTG 400							
QY	61	CGCTGCTCCCGGTCCTCGCAGGGCGCGCGCGCGCGCGCGCTTCGAGTCGGACTCG 120							
DB	401	CGTGTCTCCCGGTCCTCGCAGGGCGCGCGCGCGCGCGCGCTTCGAGTCGGACTCG 460							
QY	121	ACCTCTCGAGCGGAGCGCGAGCGCGGGCGAGCGCGCGCTTATGCAAGCAAAGATCGG 180							
DB	461	ACCTCTCGAGCGGAGCGCGAGCGCGGGCGAGCGCGCGCTTATGCAAGCAAAGATCGG 520							
QY	181	AGGAGCAGTACGCTGTGTCCAGTGTAGATGAATCATGACTGTAATCTACCCAGAT 240							
DB	521	AGGAGCAGTACGCTGTGTCCAGTGTAGATGAATCATGACTGTAATCTACCCAGAT 580							
QY	241	ATTGAAAATGTACAAGTGTCTAGCTAAGAAAGAGGCTGCGCAACATAACAGAGAACAGG 300							
DB	581	ATTGAAAATGTACAAGTGTCTAGCTAAGAAAGAGGCTGCGCAACATAACAGAGAACAGG 640							
QY	301	CCAACCTCAACTCAGGACAGAGAGACTATAAATTTGCTGCGACATTTAATAACAG 360							
DB	641	CCAACCTCAACTCAGGACAGAGAGACTATAAATTTGCTGCGACATTTAATAACAG 700							
QY	361	AGATCTTGAAGAAGTATGATATGATGAGTGAGAAAGACTCAATGCAATGCCAGGAGGTG 420							
DB	701	AGATCTTGAAGAAGTATGATATGATGAGTGAGAAAGACTCAATGCAATGCCAGGAGGTG 760							
QY	421	GATAGATGTGGGAGAGGTTTGAGTTCGGGACAAACACCTCTTTAAACCTCCATGTG 480							
DB	761	GATAGATGTGGGAGAGGTTTGAGTTCGGGACAAACACCTCTTTAAACCTCCATGTG 820							
QY	481	TGTCCGCTCACAGATGTGGGGTGTCTGCAATAGTAGAGGGCTGCAATGCAATGCAACCA 540							
DB	821	TGTCCGCTCACAGATGTGGGGTGTCTGCAATAGTAGAGGGCTGCAATGCAATGCAACCA 580							
QY	541	GCACGAGCTACCTCAGAGAGGTTATTTGAAATTAACAGTGCCTCTCTCAAGGCCCA 600							
DB	881	GCACGAGCTACCTCAGAGAGGTTATTTGAAATTAACAGTGCCTCTCTCAAGGCCCA 940							
QY	601	AACCACTAACATCAGTTTTCGCAATCAGCTTCTGCGATGATGCTTAACTGGATG 560							
DB	941	AACCACTAACATCAGTTTTCGCAATCAGCTTCTGCGATGATGCTTAACTGGATG 1000							
QY	661	TTTACAGCAAGTTTCAATCCATTTAGACGTTCCCTGCCAGCAACACTACCAAGTGC 720							
DB	1001	TTTACAGCAAGTTTCAATCCATTTAGACGTTCCCTGCCAGCAACACTACCAAGTGC 1060							
QY	721	AGCAGCGAACAGACCTGCCCAATACATGATGGAATATCAGATCTGCAGATGCC 780							
DB	1061	AGCAGCGAACAGACCTGCCCAATACATGATGGAATATCAGATCTGCAGATGCC 1120							
QY	781	TGGCTCAGAGATTTATTTTCTCGGATGCTGGAGATGACTCAACAGATGGATTC 840							
DB	1121	TGGCTCAGAGATTTATTTTCTCGGATGCTGGAGATGACTCAACAGATGGATTC 1180							
QY	841	ATGACATCTGTGACCAACAGAGGCTGATGAAGACCTGTGAGTGTGTGAGAG 900							
DB	1181	ATGACATCTGTGACCAACAGAGGCTGATGAAGACCTGTGAGTGTGTGAGAG 1240							
QY	901	CGGGGCTTCGGCTGCCAGCTGTGAGCCCCCAAGAACTAGACAGAACTCATGCGAGT 960							
DB	1241	CGGGGCTTCGGCTGCCAGCTGTGAGCCCCCAAGAACTAGACAGAACTCATGCGAGT 1300							
QY	961	GTGCTGTAAAAACAACTCTTCCCGACCAATGTGGGGCCAAACCGAGAAATTGATGAAA 1020							

Qy	1321	TGCCACAGTAGAACTGCTGTAACAGAGAGACCCCTTGTGGTCCATGCTAACAAAGACA	1380
Db	1661	TGCCACAGTAGAACTGCTGTAACAGAGAGACCCCTTGTGGTCCATGCTAACAAAGACA	1720
Qy	1381	AAAGTCTGCTTTCCCTGAACCATGTGGATTAACCTTTACAGAAATGGAGCTGAGCTCATCTG	1440
Db	1721	AAAGTCTGCTTTCCCTGAACCATGTGGATTAACCTTTACAGAAATGGAGCTGAGCTCATCTG	1780
Qy	1441	CAAAAGGCCCTTGTGTAAGACTGGTTTCTGCCAATGACCAACAGCCCAAGATTTTCCTC	1500
Db	1781	CAAAAGGCCCTTGTGTAAGACTGGTTTCTGCCAATGACCAACAGCCCAAGATTTTCCTC	1840
Qy	1501	TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAATAATTTGTTCTGC	1560
Db	1841	TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAATAATTTGTTCTGC	1900
Qy	1561	ATTCAATTTTATAGCAACAAATTTGTTAAACTCACTGTGATCAATATTTTATATCAT	1620
Db	1901	ATTCAATTTTATAGCAACAAATTTGTTAAACTCACTGTGATCAATATTTTATATCAT	1960
Qy	1621	GCAAAATATGTTTAAATAAAATGAAAAATTTGATT	1655
Db	1961	GCAAAATATGTTTAAATAAAATGAAAAATTTGATT	1995

Search completed: February 25, 2004, 19:10:37
Job time : 659 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2004, 16:56:30 ; Search time 147 Seconds
(without alignments)
5319.645 Million cell updates/sec

Title: US-09-921-143-1
Perfect score: 1674
Sequence: 1 gtcttcaccatgcactgc.....ttataaaaaaaaaaaaaa 1674

Scoring table: IDENTITY NUC
Gap 1070, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1674	100.0	1674	2	US-08-999-811-1
2	1674	100.0	1674	3	US-08-999-811-1
3	1674	100.0	1674	4	US-08-999-811-1
4	1664.6	99.4	2031	4	US-08-706-054A-1
5	1664.6	99.4	2031	4	US-08-706-054A-1
6	1663	99.3	2031	4	US-08-706-054A-2
7	1663	99.3	2031	4	US-08-706-054A-2
8	1661.2	99.2	1674	5	PTUS-096-09001-1
9	1651.8	98.7	1997	3	US-08-795-430-7
10	1651.8	98.7	1997	3	US-08-510-133A-34
11	1651.8	98.7	1997	4	US-08-355-700-7
12	1651.8	98.7	1997	4	US-08-601-132-32
13	1651.8	98.7	1997	4	US-08-671-573B-32
14	1526	91.2	1526	2	US-08-999-811-3
15	1526	91.2	1526	3	US-08-999-811-3
16	1520.2	90.8	1836	2	US-08-824-996-1
17	1192.2	70.6	1836	3	US-08-795-430-10
18	1192.2	70.6	1836	4	US-08-355-700-10
19	1182.2	70.6	1836	4	US-08-601-132-40
20	1182.2	70.6	1836	4	US-08-671-573B-40
21	1138.4	68.0	1140	3	US-08-510-133A-32
22	1138.4	68.0	1140	3	US-08-585-895-32
23	804.2	48.0	1741	3	US-08-795-430-12
24	804.2	48.0	1741	3	US-08-355-700-12
25	298	17.8	299	4	US-08-706-054A-6
26	298	17.8	299	4	US-08-313-299-6
27	144	8.6	1135	3	US-08-915-795-7

28	144	8.6	1325	3	US-08-915-795-6
29	138.4	8.3	219	3	US-08-795-430-6
30	138.4	8.3	219	3	US-08-510-133A-25
31	138.4	8.3	219	3	US-08-585-895-25
32	138.4	8.3	219	4	US-09-355-700-6
33	138.4	8.3	219	4	US-08-601-132-25
34	138.4	8.3	219	4	US-08-671-573B-25
35	130.6	7.8	2846	3	US-08-915-795-1
36	129	7.7	1337	4	US-09-016-434-982
37	129	7.7	2029	3	US-08-915-795-4
38	77	4.6	229	4	US-09-023-655-828
39	51.8	3.1	7218	1	US-08-232-463-14
40	50	3.0	50	4	US-08-706-054A-7
41	50	3.0	50	4	US-08-706-054A-8
42	50	3.0	50	4	US-09-313-299-7
43	50	3.0	50	4	US-09-313-299-8
44	48.4	2.9	495	4	US-09-244-583-25
45	48.4	2.9	627	4	US-09-244-583-27

ALIGNMENTS

RESULT 1
US-08-999-811-1
; Sequence 1, Application US/08999811
; Patent No. 5932540
; GENERAL INFORMATION:
; APPLICANT: HU, JING-SHAN
; APPLICANT: ROSEN, CRAIG A.
; APPLICANT: CAO, LIANG
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSES: STERNE, KESSLER, GOLDSTEIN & FOX
; STREET: 1100 NEW YORK AVENUE
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/999,811
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA: US 08/207,550
; FILING DATE: 8-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/465,968
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: MARKOWICZ, KAREN R.
; REGISTRATION NUMBER: 36,351
; REFERENCE/DOCKET NUMBER: 1488.1000004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)371-2600
; TELEFAX: (202)371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1674 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 12..80
; FEATURE:

Sequence 6, Appl
Sequence 6, Appl
Sequence 25, Appl
Sequence 25, Appl
Sequence 6, Appl
Sequence 25, Appl
Sequence 25, Appl
Sequence 1, Appl
Sequence 982, App
Sequence 4, Appl
Sequence 828, App
Sequence 14, Appl
Sequence 7, Appl
Sequence 8, Appl
Sequence 7, Appl
Sequence 8, Appl
Sequence 25, Appl
Sequence 27, Appl

```
; NAME/KEY: mat_peptide
; LOCATION: 81..1268
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 12..1268
; US-08-999-811-1

Query Match      100.0%; Score 1674; DB 2; Length 1674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCCTCCACCATGACCTCGCGGCTCTCTCTGTCGCGGTCTCTCTGTCGCGGCTG 60
DB 1 GTCCTCCACCATGACCTCGCGGCTCTCTCTGTCGCGGTCTCTCTGTCGCGGCTG 60
QY 61 CGCTGCTCCCGGTCTCTCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
DB 61 CGCTGCTCCCGGTCTCTCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
QY 121 ACTCTCGGACGGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
DB 121 ACTCTCGGACGGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
QY 181 AGGAGCAGTTACGGTCTGTCGTCAGTGTAGATGAATCTGACTGTACTCTACCCAGAT 240
DB 181 AGGAGCAGTTACGGTCTGTCGTCAGTGTAGATGAATCTGACTGTACTCTACCCAGAT 240
QY 241 ATTGGAATATGACAGTGTACGTTAGTAAAGGAGGCTGGCAACATTAACAGAGACAGG 300
DB 241 ATTGGAATATGACAGTGTACGTTAGTAAAGGAGGCTGGCAACATTAACAGAGACAGG 300
QY 301 CCAACTCAACTCAAGGACAGAGAGACTATATAATTTGCTCAGCACAATTAATACAG 360
DB 301 CCAACTCAACTCAAGGACAGAGAGACTATATAATTTGCTCAGCACAATTAATACAG 360
QY 361 AGATCTGAAAAGTATGATTAATGAGTGGAGAAAGACTCAATGACGACGAGGAGTGT 420
DB 361 AGATCTGAAAAGTATGATTAATGAGTGGAGAAAGACTCAATGACGACGAGGAGTGT 420
QY 421 GTATAGATGAGGAGGAGGTTGAGTGCAGACAAACACCTCTCTTAAACCTCCATGTG 480
DB 421 GTATAGATGAGGAGGAGGTTGAGTGCAGACAAACACCTCTCTTAAACCTCCATGTG 480
QY 481 TGTCCTCTACAGATGTGGGGTGTCTGCAATGAGTGGGGGCTGAGTGCATGAACACCA 540
DB 481 TGTCCTCTACAGATGTGGGGTGTCTGCAATGAGTGGGGGCTGAGTGCATGAACACCA 540
QY 541 GCAGAGCTACCTCAGCAGAGAGTTATTTGAATTTACAGTGCCTCTCTCAAGGCCCA 600
DB 541 GCAGAGCTACCTCAGCAGAGAGTTATTTGAATTTACAGTGCCTCTCTCAAGGCCCA 600
QY 601 AACCAGTAAACATCAGTTTGTGCAATCACACTTCTGCGATGCAATGCTTAACTGGATG 660
DB 601 AACCAGTAAACATCAGTTTGTGCAATCACACTTCTGCGATGCAATGCTTAACTGGATG 660
QY 661 TTACAGACAGTTTCAATTAATAGAGTTCCTGCGGAGGCTGCGGAGCAACATCAAGTGTG 720
DB 661 TTACAGACAGTTTCAATTAATAGAGTTCCTGCGGAGGCTGCGGAGCAACATCAAGTGTG 720
QY 721 AGGACGGAACAGCTCCGCCCAATTAATGATGGAATTAATCAATCTGAGATGCC 780
DB 721 AGGACGGAACAGCTCCGCCCAATTAATGATGGAATTAATCAATCTGAGATGCC 780
QY 781 TGCTCAGGAAGATTTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGATTC 840
DB 781 TGCTCAGGAAGATTTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGATTC 840
QY 841 ATGACATCTGACCAACAGGAGCTGGATGAGAGACCTGTGAGTGTCTGTCAGAG 900
DB 841 ATGACATCTGAGCAACAGGAGCTGGATGAGAGACCTGTGAGTGTCTGTCAGAG 900
QY 901 CGGGGCTTGGCCTGCGAGCTGTGGACCCCAACAAAGAACTAGACAGAACTCATGCCAGT 960
DB 901 CGGGGCTTGGCCTGCGAGCTGTGGACCCCAACAAAGAACTAGACAGAACTCATGCCAGT 960
```

```
DB 901 CGGGGCTTGGCCTGCGAGCTGTGGACCCCAACAAAGAACTAGACAGAACTCATGCCAGT 960
QY 961 GTGTCTGTAACAAACAAACTCTTCCCGAGCAATGTGGGGCCCAACCGAGAAATTTGATGAAA 1020
DB 961 GTGTCTGTAACAAACAAACTCTTCCCGAGCAATGTGGGGCCCAACCGAGAAATTTGATGAAA 1020
QY 1021 ACACATGCCAGTGTGTATGTAAAGAACTGCGCCGAGAAATCAACCCCTAAATCTGGAA 1080
DB 1021 ACACATGCCAGTGTGTATGTAAAGAACTGCGCCGAGAAATCAACCCCTAAATCTGGAA 1080
QY 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAGAAAGTTC 1140
DB 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAGAAAGTTC 1140
QY 1141 ACCACAAACATCAGCTGTACAGAGGCTGACAGAGGCTGACAGAGGCTTGTGAGC 1200
DB 1141 ACCACAAACATCAGCTGTACAGAGGCTGACAGAGGCTGACAGAGGCTTGTGAGC 1200
QY 1201 CAGGATTTTCATATAGTGAAGAGTGTGCTGCTGCTCCCTTCATATGCGCAAGACAC 1260
DB 1201 CAGGATTTTCATATAGTGAAGAGTGTGCTGCTGCTCCCTTCATATGCGCAAGACAC 1260
QY 1261 AAATGAGCTAAGATGTACTGTTTCCAGTTCATCGAATTTCTATTATGAAAACTGTGT 1320
DB 1261 AAATGAGCTAAGATGTACTGTTTCCAGTTCATCGAATTTCTATTATGAAAACTGTGT 1320
QY 1321 TGCCACAGTAGAAGTGTCTGTGAACAGAGAGACCTTGTGGTCCATGCTTAAACAGACA 1380
DB 1321 TGCCACAGTAGAAGTGTCTGTGAACAGAGAGACCTTGTGGTCCATGCTTAAACAGACA 1380
QY 1381 AAAAGTCTGCTTTTCTGAAACCTGTCGATAACTTTACAGAAATGGAGCTGAGCTCATCTG 1440
DB 1381 AAAAGTCTGCTTTTCTGAAACCTGTCGATAACTTTACAGAAATGGAGCTGAGCTCATCTG 1440
QY 1441 CAAAAGGCTCTTGTAAAGACTGGTTTTCTGCCAATGACCAACACGACCCAGATTTTCCCTC 1500
DB 1441 CAAAAGGCTCTTGTAAAGACTGGTTTTCTGCCAATGACCAACACGACCCAGATTTTCCCTC 1500
QY 1501 TTGTGATTTCTTTAAAGAAATGACTATAAATTTATTTCCACTTAAATAATTTTCTGTC 1560
DB 1501 TTGTGATTTCTTTAAAGAAATGACTATAAATTTATTTCCACTTAAATAATTTTCTGTC 1560
QY 1561 ATTCTATTTTATAGCAACAACATTTGGTAAACTCTGATGATCAATTTTATATCAT 1620
DB 1561 ATTCTATTTTATAGCAACAACATTTGGTAAACTCTGATGATCAATTTTATATCAT 1620
QY 1621 GCAAAATATCTTTAAATAAAATGAAAATGTAATTTATATAAAAAA 1674
DB 1621 GCAAAATATCTTTAAATAAAATGAAAATGTAATTTATATAAAAAA 1674
```

RESULT 2

```
US-09-042-105-1
; Sequence 1, Application US/09042105
; Patent No. 6040157
; GENERAL INFORMATION:
; APPLICANT: HU, JING-SHAN
; APPLICANT: ROSEN, CRAIG A.
; APPLICANT: CAO, LIANG
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
; STREET: 1100 NEW YORK AVENUE
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/042,105
 ; FILING DATE: HEREWITH
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/207,550
 ; FILING DATE: 8-MAR-1994
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/465,968
 ; FILING DATE: 06-JUN-1995
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: TO BE ASSIGNED
 ; FILING DATE: 24-DEC-1997
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ERIC K. STEFFE
 ; REGISTRATION NUMBER: 36,698
 ; REFERENCE/DOCKET NUMBER: 1488.1000003/EKS
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202)371-2600
 ; TELEFAX: (202)371-2540
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1674 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cdna
 ; FEATURE:
 ; NAME/KEY: sig_peptide
 ; LOCATION: 12..80
 ; FEATURE:
 ; NAME/KEY: mat_peptide
 ; LOCATION: 81..1268
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 12..1268
 ;

US-09-042-105-1

Query Match 100.0%; Score 1674; DB 3; Length 1674;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GTCCTTCCACCAATGCACTCGCTGGCTTCTCTCTGTGGGGTTCCTCTGCTCGCGCTG	60
Db	1	GTCCTTCCACCAATGCACTCGCTGGCTTCTCTCTGTGGGGTTCCTCTGCTCGCGCTG	60
Qy	61	CGCTGCTCCGGGTCCTCGGAGCGCCCGCGCGCGCTTCGAGTCGGACTCG	120
Db	61	CGCTGCTCCGGGTCCTCGGAGCGCCCGCGCGCGCTTCGAGTCGGACTCG	120
Qy	121	ACCTCTCGGAGCGGAGCCCGAGCGCGGAGCCACCGCTTATGCAAGCAAGATCTGG	180
Db	121	ACCTCTCGGAGCGGAGCCCGAGCGCGGAGCCACCGCTTATGCAAGCAAGATCTGG	180
Qy	181	AGGAGCAGTACGCTGTGTCAGTGTAGATCACTGACTGACTCTACCCAGAT	240
Db	181	AGGAGCAGTACGCTGTGTCAGTGTAGATCACTGACTGACTCTACCCAGAT	240
Qy	241	ATTGGAATATGACAGTGTCAAGTGTAGAGGAGGCTGGCAACATACAGAACAGG	300
Db	241	ATTGGAATATGACAGTGTCAAGTGTAGAGGAGGCTGGCAACATACAGAACAGG	300
Qy	301	CCAACTCACTCAAGGACGAGAGACTATAAATTTGCTGAGCAGACATTAATACAG	360
Db	301	CCAACTCACTCAAGGACGAGAGACTATAAATTTGCTGAGCAGACATTAATACAG	360
Qy	361	AGATCTTGAAAGATTTGATAATGAGTGGAGAAAGACTCAATGATGCCACGGAGTGT	420
Db	361	AGATCTTGAAAGATTTGATAATGAGTGGAGAAAGACTCAATGATGCCACGGAGTGT	420

Qy	421	GTATAGATGGGGAAGAGTTTGGAGTCGCGACAAACACCTTCTTAAACCTCCATGTG	480
Db	421	GTATAGATGGGGAAGAGTTTGGAGTCGCGACAAACACCTTCTTAAACCTCCATGTG	480
Qy	481	TGTCCTGTACAGATGTGGGGTGTGCAATAGTAGAGGGGTGCGATGCATGAACACA	540
Db	481	TGTCCTGTACAGATGTGGGGTGTGCAATAGTAGAGGGGTGCGATGCATGAACACA	540
Qy	541	GCAGAGTACCTCAGCAGAGCGTTATTTGAATTTACAGTGCCTCTCTCAAGGCCCA	600
Db	541	GCAGAGTACCTCAGCAGAGCGTTATTTGAATTTACAGTGCCTCTCTCAAGGCCCA	600
Qy	601	AACAGTAAACATCAGTTTGGCAATCACTCTCCGCGGATGCATGTCTAACTGGATG	660
Db	601	AACAGTAAACATCAGTTTGGCAATCACTCTCCGCGGATGCATGTCTAACTGGATG	660
Qy	661	TTTACAGACAAAGTTTCAATTCATTTAGACGTTTCCCTGCCAGCAACACTTACCAGTGTG	720
Db	661	TTTACAGACAAAGTTTCAATTCATTTAGACGTTTCCCTGCCAGCAACACTTACCAGTGTG	720
Qy	721	AGGAGCGAAACAGACCTGCCCAACCAATTAATGTGGAATTAATCAATCTCAGATGCC	780
Db	721	AGGAGCGAAACAGACCTGCCCAACCAATTAATGTGGAATTAATCAATCTCAGATGCC	780
Qy	781	TGGCTCAGGAAGATTTTATGTTTCTCGGATGCTGGAGTCACTCAACAGATGATGCC	840
Db	781	TGGCTCAGGAAGATTTTATGTTTCTCGGATGCTGGAGTCACTCAACAGATGATGCC	840
Qy	841	ATGACATCTGTGGACCAACAAAGAGCTGGATGAAGAGACCTGTGAGTGTCTCAGAG	900
Db	841	ATGACATCTGTGGACCAACAAAGAGCTGGATGAAGAGACCTGTGAGTGTCTCAGAG	900
Qy	901	CGGGCTTCGGCTGCGAGCTGTGGACCCCAACAAAGAACTAGACAGAACTCATGCCAGT	960
Db	901	CGGGCTTCGGCTGCGAGCTGTGGACCCCAACAAAGAACTAGACAGAACTCATGCCAGT	960
Qy	961	GTGCTGTAAAAACAAACTCTTCCCGAGCAATGTGGGCGCAACCGAGATTTGATGAAA	1020
Db	961	GTGCTGTAAAAACAAACTCTTCCCGAGCAATGTGGGCGCAACCGAGATTTGATGAAA	1020
Qy	1021	ACATGCGCAGTGTATGTAAAAAGAACTGCCCGCAAGAACTCAACCTTAACTCTGAAA	1080
Db	1021	ACATGCGCAGTGTATGTAAAAAGAACTGCCCGCAAGAACTCAACCTTAACTCTGAAA	1080
Qy	1081	AATGTGCTGTGAATGTACAGAAAGTCCAGAAATGCTTGTAAAAAGAAAGATGCC	1140
Db	1081	AATGTGCTGTGAATGTACAGAAAGTCCAGAAATGCTTGTAAAAAGAAAGATGCC	1140
Qy	1141	ACCACCAACATGCAGCTGTACAGACGGCCATGTACGAAACCGCCAGAGGCTTGTGAGC	1200
Db	1141	ACCACCAACATGCAGCTGTACAGACGGCCATGTACGAAACCGCCAGAGGCTTGTGAGC	1200
Qy	1201	CAGGATTTTCAATAGTGAAGAGTGTGCTGTGTCCTTCATATTGGGAAAGACAC	1260
Db	1201	CAGGATTTTCAATAGTGAAGAGTGTGCTGTGTCCTTCATATTGGGAAAGACAC	1260
Qy	1261	AAATGAGTAAAGATTTCTACTTTTCCAGTTTCACTGATTTTCTATTGGAAAACTGTGT	1320
Db	1261	AAATGAGTAAAGATTTCTACTTTTCCAGTTTCACTGATTTTCTATTGGAAAACTGTGT	1320
Qy	1321	TGCCACAGTAGAATCTGTCTGTGAACAGAGACCCCTTGGGGTCCATGCTAAACAAGACA	1380
Db	1321	TGCCACAGTAGAATCTGTCTGTGAACAGAGACCCCTTGGGGTCCATGCTAAACAAGACA	1380
Qy	1381	AAAGTGTCTTTTCCCTGAACCATGTGATACTTTACAGAAATGGACTGGAGCTCATCTG	1440
Db	1381	AAAGTGTCTTTTCCCTGAACCATGTGATACTTTACAGAAATGGACTGGAGCTCATCTG	1440
Qy	1441	CAAAAGCCCTCTGTAAAGACTGGTTTCTGCAATGACCAACAGCCAGAGATTTTCTCTC	1500
Db	1441	CAAAAGCCCTCTGTAAAGACTGGTTTCTGCAATGACCAACAGCCAGAGATTTTCTCTC	1500
Qy	1501	TTGTGATTTTCTTTAAAGAAATGACTATATAATTTTTCCTCACTAAAAATATTGTTCTGC	1560

```
Db 1501 TTGTGATTTCTTAAAGAAATGACATATATATATTTATTTCCACTAAAAATATTTGTTCTGC 1560
Qy 1561 ATTCAATTTTATAGCAACAAATGGTAAACTCAGTGTGATCAATATTTTATATCAT 1620
Db 1561 ATTCAATTTTATAGCAACAAATGGTAAACTCAGTGTGATCAATATTTTATATCAT 1620
Qy 1621 GCAAAATATGTTTAAATAAATGAAATTTGATTTATATAAAAAA 1674
Db 1621 GCAAAATATGTTTAAATAAATGAAATTTGATTTATATAAAAAA 1674

RESULT 3
US-08-465-968-1
; Sequence 1, Application US/08465968E
; Patent No. 6608182
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Vascular Endothelial Growth Factor 2
; FILE REFERENCE: PF112PI
; CURRENT APPLICATION NUMBER: US/08/465,968E
; CURRENT FILING DATE: 1995-06-06
; EARLIER APPLICATION NUMBER: 08/207,550
; EARLIER FILING DATE: 1994-03-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1674
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (12)..(1268)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (12)..(149)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (150)..(1268)
US-08-465-968-1

Query Match 100.0%; Score 1674; DB 4; Length 1674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTCTTCCACCATGCACTCGCTGGGCTTCTTCTGTGGCGGTTCCTCTGCTGCGCGCTG 60
Db 1 GTCTTCCACCATGCACTCGCTGGGCTTCTTCTGTGGCGGTTCCTCTGCTGCGCGCTG 60

Qy 61 CGTGTCTCCGGGTCTCTCGAGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCG 120
Db 61 CGTGTCTCCGGGTCTCTCGAGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCG 120

Qy 121 ACCTCTCGAGCGGAGCGCGAGCGGGCGAGCCAGCGCTTATGCAAGCAAGATCTGG 180
Db 121 ACCTCTCGAGCGGAGCGCGAGCGGGCGAGCCAGCGCTTATGCAAGCAAGATCTGG 180

Qy 181 AGGAGCAGTTACGGTCTGTGTCCAGTGTAGATGAATCATGACTGTACTTACCCAGAAT 240
Db 181 AGGAGCAGTTACGGTCTGTGTCCAGTGTAGATGAATCATGACTGTACTTACCCAGAAT 240

Qy 241 ATTGGAATATGACAGTGTACAGTAAAGAGAGGCTGGCAACATAACAGAGACAGG 300
Db 241 ATTGGAATATGACAGTGTACAGTAAAGAGAGGCTGGCAACATAACAGAGACAGG 300

Qy 301 CCAACTCACTCAAGCAGAGAGACTATAAAATTTGCTGAGCAGACATATAATACAG 360
Db 301 CCAACTCACTCAAGCAGAGAGACTATAAAATTTGCTGAGCAGACATATAATACAG 360

Qy 361 AGATCTTGAAGATATGATTAATAGTGGAGAAAGACTCAATGCAATGCCAGGGAGGTG 420
Db 361 AGATCTTGAAGATATGATTAATAGTGGAGAAAGACTCAATGCAATGCCAGGGAGGTG 420
```

```
Qy 421 GTATAGATCTGGGAAGAGATTGGAGTGGCGACAAAACACCTTCTTTAAACCTCATGTG 480
Db 421 GTATAGATCTGGGAAGAGATTGGAGTGGCGACAAAACACCTTCTTTAAACCTCATGTG 480

Qy 481 TGTCCGCTCTACAGATGTGGGGGTGTGTGCAATAGTAGAGGGGTGAGTGCATGAACCA 540
Db 481 TGTCCGCTCTACAGATGTGGGGGTGTGTGCAATAGTAGAGGGGTGAGTGCATGAACCA 540

Qy 541 GACGAGCTACCTCAGCAAGACGTTATTTGAAATTTACAGTGGCTCTCTCTCAAGGCCCA 600
Db 541 GACGAGCTACCTCAGCAAGACGTTATTTGAAATTTACAGTGGCTCTCTCTCAAGGCCCA 600

Qy 601 AACCACTAACATCAGTTTGGCAATCAACATTTCTGCGGATGCATGTCTAAACTGGATG 660
Db 601 AACCACTAACATCAGTTTGGCAATCAACATTTCTGCGGATGCATGTCTAAACTGGATG 660

Qy 661 TTTACAGACAAGTTCAATTCATTTAGACGTTCCCTGCCAGCAACACTACCACAGTGC 720
Db 661 TTTACAGACAAGTTCAATTCATTTAGACGTTCCCTGCCAGCAACACTACCACAGTGC 720

Qy 721 AGGCGGCAACAGACCTGCCCAACCAATTACATGTGGAATATACATCTGCAGATGCC 780
Db 721 AGGCGGCAACAGACCTGCCCAACCAATTACATGTGGAATATACATCTGCAGATGCC 780

Qy 781 TGGCTCAGGAAGATTTTATGTTTCTCGATGCTGGAGATGACTCAACAGATGGATCC 840
Db 781 TGGCTCAGGAAGATTTTATGTTTCTCGATGCTGGAGATGACTCAACAGATGGATCC 840

Qy 841 ATGACATCTGTGGACCAAAACAGAGCTGGATGAAGAGACCTGTCACTGTCTGCAGAG 900
Db 841 ATGACATCTGTGGACCAAAACAGAGCTGGATGAAGAGACCTGTCACTGTCTGCAGAG 900

Qy 901 CGGGGCTTCCGGCTGCGAGCTGTGACACCCCAAAAGAACTAGACAGAAACTCATGCCAGT 960
Db 901 CGGGGCTTCCGGCTGCGAGCTGTGACACCCCAAAAGAACTAGACAGAAACTCATGCCAGT 960

Qy 961 GTGTCTGTAAAAACAACTCTTCCAGCCCAATGTGGGGCCAAACCGAGAAATTTGATGAA 1020
Db 961 GTGTCTGTAAAAACAACTCTTCCAGCCCAATGTGGGGCCAAACCGAGAAATTTGATGAA 1020

Qy 1021 ACACATGCCAGTGTGTATGTAAAGAACTGTCGCCCAAGAAATCAACCCCTAAATCTCGAA 1080
Db 1021 ACACATGCCAGTGTGTATGTAAAGAACTGTCGCCCAAGAAATCAACCCCTAAATCTCGAA 1080

Qy 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTGTGTAAAGAAAGAGTTCC 1140
Db 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTGTGTAAAGAAAGAGTTCC 1140

Qy 1141 ACCACCAACATCAGCTGTTACAGACGGCCATGTACGAAACCGCCAGAGGCTTGTGAGC 1200
Db 1141 ACCACCAACATCAGCTGTTACAGACGGCCATGTACGAAACCGCCAGAGGCTTGTGAGC 1200

Qy 1201 CAGGATTTTCATATAGTGAAGAGTGTGTGTTGTCTCCCTTCATATTTGGCAAGACCAC 1260
Db 1201 CAGGATTTTCATATAGTGAAGAGTGTGTGTTGTCTCCCTTCATATTTGGCAAGACCAC 1260

Qy 1261 AATGAGCTAAGATTGACTGTTTTCCAGTTCATCGATTTCTTATTTATGAAACTGTGT 1320
Db 1261 AATGAGCTAAGATTGACTGTTTTCCAGTTCATCGATTTCTTATTTATGAAACTGTGT 1320

Qy 1321 TGCCACAGTGAAGACTGTCTGTAACAGAGAGACCCCTTGTGGGTCCATGTCTAAACAGACA 1380
Db 1321 TGCCACAGTGAAGACTGTCTGTAACAGAGAGACCCCTTGTGGGTCCATGTCTAAACAGACA 1380

Qy 1381 AAGTCTGTCTTCTGCAACATGTGGATAAATTTACAGAAATGGAGTGGAGTCTATCTG 1440
Db 1381 AAGTCTGTCTTCTGCAACATGTGGATAAATTTACAGAAATGGAGTGGAGTCTATCTG 1440

Qy 1441 CAAAGGCCCTCTGTAAGAGACTGTTTTCTGCCAATGACCCAAACAGCCCAAGATTTTCTC 1500
Db 1441 CAAAGGCCCTCTGTAAGAGACTGTTTTCTGCCAATGACCCAAACAGCCCAAGATTTTCTC 1500

Qy 1501 TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAAAATATTTGTTCTGC 1560
```


Db 1501 TTGTGATTTCTTTAAAGAAATGACTATATATATTTTTCACATAAAATAATTTCTGTC 1560
 Qy 1561 ATTCAATTTTATAGCAACAACAAATTTGTAATAAATCTCACTGTGATCAATATTTTATATCAT 1620
 Db 1561 ATTCAATTTTATAGCAACAACAAATTTGTAATAAATCTCACTGTGATCAATATTTTATATCAT 1620
 Qy 1561 GCAAAATATGTTTAAATAAATAAATAAATTTGTAATTTTAAATAAATAAATAAATAA 1674
 Db 1561 GCAAAATATGTTTAAATAAATAAATAAATTTGTAATTTTAAATAAATAAATAAATAA 1674

RESULT 4

US-08-706-054A-1
 ; Sequence 1, Application US/08705054A
 ; Patent No. 6451764
 ; GENERAL INFORMATION:
 ; APPLICANT: Lee, James
 ; Wood, William I.
 ; TITLE OF INVENTION: VEGF-Related Protein
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genentech, Inc.
 ; STREET: 460 Point San Bruno Blvd
 ; CITY: South San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94080

COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: WinPatIn (Genentech)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/706,054A
 ; FILING DATE: 30-Aug-1996
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/003491
 ; FILING DATE: 08-Sep-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lee, Wendy M.
 ; REGISTRATION NUMBER: P-40,378
 ; REFERENCE/DOCKET NUMBER: P0963R1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415/225-1994
 ; TELEFAX: 415/952-9881
 ; TELEX: 910/371-7168
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2031 base pairs
 ; TYPE: Nucleic Acid
 ; STRANDEDNESS: Single
 ; TOPOLOGY: Linear
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-08-706-054A-1

Query Match 99.4%; Score 1664.6; DB 4; Length 2031;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1667; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 1 GTCTTCCACCATGACATCGTGGGCTTCTCTGTGGGCTGTCTCTGCTCGCGCTG 60
 Db 361 GTCTTCCACCATGACATCGTGGGCTTCTCTGTGGGCTGTCTCTGCTCGCGCTG 420
 Qy 61 CGCTGCTCCCGGTCCTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
 Db 421 CGCTGCTCCCGGTCCTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
 Qy 121 ACCTCTCGGACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
 Db 481 ACCTCTCGGACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540

Qy 181 AGGACAGTTACGCTCTGTCTCCAGTGTAGATGAATCACTCATGACTGTACTCTACCCAGAT 240
 Db 541 AGGACAGTTACGCTCTGTCTCCAGTGTAGATGAATCACTCATGACTGTACTCTACCCAGAT 600
 Qy 241 ATTGGAATAATGTACAAGTGTACAGTGTAGAGAAAGAGGCTGGCAACATACAGAGAACAGG 300
 Db 601 ATTGGAATAATGTACAGTGTACAGTGTAGAGAAAGAGGCTGGCAACATACAGAGAACAGG 660
 Qy 301 CCAACCTCAACTCAAGGACAGAGACTATAAAATTTGCTGTGAGCAGCATTAATATACAG 360
 Db 661 CCAACCTCAACTCAAGGACAGAGACTATAAAATTTGCTGTGAGCAGCATTAATATACAG 720
 Qy 361 AGATCTTGAAAGTATTGATATGATGAGAGAGACTCAATGATGATGATGATGATGATGATG 420
 Db 721 AGATCTTGAAAGTATTGATATGATGATGATGATGATGATGATGATGATGATGATGATG 780
 Qy 421 GTATAGATGTGGGAAAGAGTTTGGAGTGGCAAAACACCTTTTAAACCTTCATGTG 480
 Db 781 GTATAGATGTGGGAAAGAGTTTGGAGTGGCAAAACACCTTTTAAACCTTCATGTG 840
 Qy 481 TGTCCGCTACAGATGTGGGAGTTGCTGCAATAGTGGGGCTGCAGTGCATGACACCA 540
 Db 841 TGTCCGCTACAGATGTGGGAGTTGCTGCAATAGTGGGGCTGCAGTGCATGACACCA 900
 Qy 541 GCACGAGTACTCTCAGCAAGACGTTATTGAAATTTACAGTGCCTCTCTCTCAAGSCCCA 600
 Db 901 GCACGAGTACTCTCAGCAAGACGTTATTGAAATTTACAGTGCCTCTCTCTCAAGSCCCA 960
 Qy 601 AACAGTAACAAATGATTTGCGCAATCACTTCTGCGGATGATGATGATGATGATGATG 660
 Db 961 AACAGTAACAAATGATTTGCGCAATCACTTCTGCGGATGATGATGATGATGATGATG 1020
 Qy 661 TTTACAGACAAGTTTCATTTCCATTATTAGACGTTCCCTGCCAGCACTACACAGTGC 720
 Db 1021 TTTACAGACAAGTTTCATTTCCATTATTAGACGTTCCCTGCCAGCACTACACAGTGC 1080
 Qy 721 AGGACAGCAACAGACCTGCCCCCACTTATGATGATGATGATGATGATGATGATGATG 780
 Db 1081 AGGACAGCAACAGACCTGCCCCCACTTATGATGATGATGATGATGATGATGATGATG 1140
 Qy 781 TGGCTCAGGAGATTTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGATGATG 840
 Db 1141 TGGCTCAGGAGATTTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGATGATG 1200
 Qy 841 ATGACATCTGTGGACCAACAGAGCTGGATGAAGAGACCTGTCACTGTCTGTGAGAG 900
 Db 1201 ATGACATCTGTGGACCAACAGAGCTGGATGAAGAGACCTGTCACTGTCTGTGAGAG 1260
 Qy 901 CGGGCTTCGGCTCGGCTCGGCTGTGGACCCCAAGAGACTAGACAGAACTCATGCCAGT 960
 Db 1261 CGGGCTTCGGCTCGGCTCGGCTGTGGACCCCAAGAGACTAGACAGAACTCATGCCAGT 1320
 Qy 961 GTGCTGTAAAAACAAACTCTTCCCGAGCAATGTGGGGCCCAACCGAGAAATTTGATGAA 1020
 Db 1321 GTGCTGTAAAAACAAACTCTTCCCGAGCAATGTGGGGCCCAACCGAGAAATTTGATGAA 1380
 Qy 1021 ACATATGCGAGTGTATGTAAAGAACTGCGCCAGAAATCAACCCCTTAATCTCGAA 1080
 Db 1381 ACATATGCGAGTGTATGTAAAGAACTGCGCCAGAAATCAACCCCTTAATCTCGAA 1440
 Qy 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAGAGTTCC 1140
 Db 1441 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAGAGTTCC 1500
 Qy 1141 ACCACCAACATGACGTGTTTACAGACCGCATGTAACGCGCAGAGGCTTGTGAGC 1200
 Db 1501 ACCACCAACATGACGTGTTTACAGACCGCATGTAACGCGCAGAGGCTTGTGAGC 1560
 Qy 1201 CAGGATTTTCATATAGTGAAGAGTGTGCTGTGCTCCCTTCATTTGGGAAAGACAC 1260
 Db 1561 CAGGATTTTCATATAGTGAAGAGTGTGCTGTGCTCCCTTCATTTGGGAAAGACAC 1620
 Qy 1261 AAATGAGCTAAGATTTGTACTGTGTTTCCAGTTTCATGATTTTCTATATGGAATGTTGT 1320

1621	DB	AAATGAGCTAAGATTGCTACTGTTTCCAGTTTCATCGATTTCTATTATGAAAACTGTGT	1680
1321	QY	TGCCACAGTAAAGACTGCTGTGAACAGAGAGACCCTTGCGGTCCATGCTAAACAAAGACA	1380
1681	DB	TGCCACAGTAAAGACTGCTGTGAACAGAGAGACCCTTGCGGTCACTGCTAACAAAGACA	1740
1381	QY	AAAGTCGTCTTTTCTTCTGAACCACTGTGGATAACTTTACAGAAATGGACTGGAGCTCATCTG	1440
1741	DB	AAAGTCGTCTTTTCTTCTGAACCACTGTGGATAACTTTACAGAAATGGACTGGAGCTCATCTG	1800
1441	QY	CAAAAGGCCCTCTGTAAAGAGACTGGTTTTCTGCGCAATGACCAACAGACCAAGATTTTCCCTC	1500
1801	DB	CAAAAGGCCCTCTGTAAAGAGACTGGTTTTCTGCGCAATGACCAACAGACCAAGATTTTCCCTC	1860
1501	QY	TTGTGATTTCTTTTAAAGAAATGACTATATAATTTATTTCCACTAAAAATATGTTTCTGCG	1560
1861	DB	TTGTGATTTCTTTTAAAGAAATGACTATATAATTTATTTCCACTAAAAATATGTTTCTGCG	1920
1561	QY	ATTTCATTTTATATGCAACACAAATTCGGTAAAACTCAGCTGTGATCAATATTTTATATCAT	1620
1921	DB	ATTTCATTTTATATGCAACACAAATTCGGTAAAACTCAGCTGTGATCAATATTTTATATCAT	1980
1621	QY	GCAAAATATGTTTAAAAATAAAATGAAAAATGTTATTTATAAAAAAATAAAAAA	1671
1981	DB	GCAAAATATGTTTAAAAATAAAATGAAAAATGTTATTTAAAAAATAAAAAAATAAAAAA	2031

1321 TGGCAGTAGAAGTCTGTGTGAAACAGAGAGACCCCTTGTGGTCCATCTAACAAAGACA 1380
 Db TGGCAGTAGAAGTCTGTGTGAAACAGAGAGACCCCTTGTGGTCCATCTAACAAAGACA 1740
 1381 AAGAGTGTCTTCTTCAACCAATGTGGATACTTTACAGAAATGAGCTGGAGCTCATCTG 1440
 Db AAGAGTGTCTTCTTCAACCAATGTGGATACTTTACAGAAATGAGCTGGAGCTCATCTG 1800
 1441 CAAAAGCCCTCTGTAAAGAGCTGTTTTCTGCAATGACCAACAGCCAAAGATTTCTTC 1500
 Db CAAAAGCCCTCTGTAAAGAGCTGTTTTCTGCAATGACCAACAGCCAAAGATTTCTTC 1860
 1501 TTGTGATTTCTTTAAAGAGTACATATAATTTTCCACTAAAAATATTGTTCTGTC 1560
 Db TTGTGATTTCTTTAAAGAGTACATATAATTTTCCACTAAAAATATTGTTCTGTC 1920
 1561 ATTCAATTTTATAGCAACAACTTGTAAAGAGTACATATAATTTTCCACTAAAAATATTGTTCTGTC 1620
 Db ATTCAATTTTATAGCAACAACTTGTAAAGAGTACATATAATTTTCCACTAAAAATATTGTTCTGTC 1980
 1621 GCAAAATATGTTTAAAGATGAAATGAAATTTGTTTATAAAAAA 1671
 Db GCAAAATATGTTTAAAGATGAAATGAAATTTGTTTATAAAAAA 2031

RESULT 6

US-08-706-054A-2/c

; Sequence 2, Application US/08706054A

; Patent No. 6451764

; GENERAL INFORMATION:

; APPLICANT: Lee, James

; TITLE OF INVENTION: VEGF-Related Protein

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 460 Point San Bruno Blvd

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WinPatIn (Genentech)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/706, 054A

; FILING DATE: 30-Aug-1996

; CLASSIFICATION: <Unknown>

; PRIORITY DATA:

; APPLICATION NUMBER: 60/003491

; FILING DATE: 08-Sep-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Lee, Wendy M.

; REGISTRATION NUMBER: P-40, 378

; REFERENCE/DOCKET NUMBER: P0963R1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415/225-1994

; TELEFAX: 415/952-9881

; TELEX: 910/371-7168

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2031 base pairs

; TYPE: Nucleic Acid

; STRANDEDNESS: Single

; TOPOLOGY: Linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

; US-08-706-054A-2

; Query Match

; Best Local Similarity

; 99.3%; Score 1663; DB 4; Length 2031;

; 99.7%; Pred. No. 0;

Matches 1666; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Qy 1 GTCTTCCACCATGCACTCGTGGGCTTCTTCTGTGGGCTGTCTCTGTCCCGCTG 60
 Db 1671 GTCTTCCACCATGCACTCGTGGGCTTCTTCTGTGGGCTGTCTCTGTCCCGCTG 1612
 Qy 61 CGCTCTCCCGGCTCTCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
 Db 1611 TGCTCTCCCGGCTCTCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1552
 Qy 121 ACCTTCGAGACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
 Db 1551 ACCTTCGAGACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1492
 Qy 181 AGGACGAGTTAGGCTCTGTGTCAGTGTAGTGAACCTCATGACTGTACTCTACCCAGAT 240
 Db 1491 AGGACGAGTTAGGCTCTGTGTCAGTGTAGTGAACCTCATGACTGTACTCTACCCAGAT 1432
 Qy 241 ATTGAAAAATGTACAAGTGTGAGCTTAAGGAAAGGAGGCTGGCAACATAACAGAGAAC 300
 Db 1431 ATTGAAAAATGTACAAGTGTGAGCTTAAGGAAAGGAGGCTGGCAACATAACAGAGAAC 1372
 Qy 301 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGAGCAGCATTTATAATACAG 360
 Db 1371 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGAGCAGCATTTATAATACAG 1312
 Qy 361 AGATCTTGAAGATTTGATTAATGAGTGGAGAAAGACTCAATGATGCCACGGAGGTGT 420
 Db 1311 AGATCTTGAAGATTTGATTAATGAGTGGAGAAAGACTCAATGATGCCACGGAGGTGT 1252
 Qy 421 GTATAGATGTGGGAAAGGAGTTTGGAGTCGCGCAAAACACCTTCTTAAACCTCCATGTG 480
 Db 1251 GTATAGATGTGGGAAAGGAGTTTGGAGTCGCGCAAAACACCTTCTTAAACCTCCATGTG 1192
 Qy 481 TGTCCGTCTACAGATGTGGGGTTGCTGCAATAGTGGAGGGCTGCATGATGAACACCA 540
 Db 1191 TGTCCGTCTACAGATGTGGGGTTGCTGCAATAGTGGAGGGCTGCATGATGAACACCA 1132
 Qy 541 GCACGAGTCTACTCAGCAAGAGCTTTATTGAAATTTACAGTGCCCTCTCTCAAGGCCCA 600
 Db 1131 GCACGAGTCTACTCAGCAAGAGCTTTATTGAAATTTACAGTGCCCTCTCTCAAGGCCCA 1072
 Qy 601 AACCAAGTAAATCAGTTTGGCAATCAGCTTCTCCGCGATGATGTCTAAACGTGATG 660
 Db 1071 AACCAAGTAAATCAGTTTGGCAATCAGCTTCTCCGCGATGATGTCTAAACGTGATG 1012
 Qy 661 TTACAGACAAAGTTCATTCCATTTAGACGCTTCCCTGCCAGCAACACTACCAAGTGT 720
 Db 1011 TTACAGACAAAGTTCATTCCATTTAGACGCTTCCCTGCCAGCAACACTACCAAGTGT 952
 Qy 721 AGGACGAGAAACAGACCTGCCCCCAATTAATGAGTGAATTAATCATCTGAGATGCC 780
 Db 951 AGGACGAGAAACAGACCTGCCCCCAATTAATGAGTGAATTAATCATCTGAGATGCC 892
 Qy 781 TGGCTCAGAAAGATTTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGGATTCC 840
 Db 891 TGGCTCAGAAAGATTTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGGATTCC 832
 Qy 841 ATGACATCTGTGGACCAAAACAGGAGTGTGATGAAGAGACCTGTGAGTGTCTGAGAG 900
 Db 831 ATGACATCTGTGGACCAAAACAGGAGTGTGATGAAGAGACCTGTGAGTGTCTGAGAG 772
 Qy 901 CGGGGCTTCCGGCTGCCAGCTGTGAGCCCAAGAAAGAACTAGACAGAAACTCATGCCAGT 960
 Db 771 CGGGGCTTCCGGCTGCCAGCTGTGAGCCCAAGAAAGAACTAGACAGAAACTCATGCCAGT 712
 Qy 961 GTGTCTGTAAAAAACAACCTTTCCCGAGCCATGTGGGGCCACCGAGATTTGATGAAA 1020
 Db 711 GTGTCTGTAAAAAACAACCTTTCCCGAGCCAACTGTGGGGCCCAACCGAGATTTGATGAAA 652
 Qy 1021 ACACATGCCAGTGTGTATGTAAGAAACCTTGGCCCGAGAAATCAACCCCTAAATCCCTGGAA 1080
 Db 651 ACACATGCCAGTGTGTATGTAAGAAACCTTGGCCCGAGAAATCAACCCCTAAATCCCTGGAA 592

QY 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTGTTAAAGGAAAGAGTTCC 1140
 Db 591 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTGTTAAAGGAAAGAGTTCC 532
 QY 1141 ACCACCAAAACATGACGCTGTATACAGACGGCCATGTACAGACCGCCAGAGGCTTGTGAGC 1200
 Db 531 ACCACCAAAACATGACGCTGTATACAGACGGCCATGTACAGACCGCCAGAGGCTTGTGAGC 472
 QY 1201 CAGGATTTTCAATAGTAGAAGAGTGTGCTGTGCTGCTTCAATATGCGAAAGACAC 1260
 Db 471 CAGGATTTTCAATAGTAGAAGAGTGTGCTGTGCTGCTTCAATATGCGAAAGACAC 412
 QY 1261 AAATGAGCTAAGATTGTACCTGTTTCCAGTTTCATCGATTTCTATTATGGAATACTGTGT 1320
 Db 411 AAATGAGCTAAGATTGTACCTGTTTCCAGTTTCATCGATTTCTATTATGGAATACTGTGT 352
 QY 1321 TGCCACAGTAGAAGTGTGCTGTGACAGAGAGACCCCTTGTGGTCCATGCTAAACAAGACA 1380
 Db 351 TGCCACAGTAGAAGTGTGCTGTGACAGAGAGACCCCTTGTGGTCCATGCTAAACAAGACA 292
 QY 1381 AAAGTCTGTCTTCTGAAACCATGTGATGAACTTTACAGAAATGACTGGAGCTCATCTG 1440
 Db 291 AAAGTCTGTCTTCTGAAACCATGTGATGAACTTTACAGAAATGACTGGAGCTCATCTG 232
 QY 1441 CAAAGGCTCTTGTAAAGACTGGTTTCTGCCAATGACCAACAGCCAGATTTTCTC 1500
 Db 231 CAAAGGCTCTTGTAAAGACTGGTTTCTGCCAATGACCAACAGCCAGATTTTCTC 172
 QY 1501 TTGATGATTTCTTAAAGAAAGTACTATATAATTTATTTCCACTAAAAATATTGTTTCTGC 1560
 Db 171 TTGATGATTTCTTAAAGAAAGTACTATATAATTTATTTCCACTAAAAATATTGTTTCTGC 112
 QY 1561 ATTCAATTTTATAGCAACACATTTGGTAAACTGACTGTGATCAATATTTTATATCAT 1620
 Db 111 ATTCAATTTTATAGCAACACATTTGGTAAACTGACTGTGATCAATATTTTATATCAT 52
 QY 1621 GCAAAATATGTTTAAATATAAAATGAAAAATGTTATTTATAAAAAATAAAAAA 1671
 Db 51 GCAAAATATGTTTAAATATAAAATGAAAAATGTTATTTATAAAAAATAAAAAA 1

RESULT 7

US-09-313-299-2/c
 ; Sequence 2, Application US/09313299B
 ; Patent No. 6576608
 ; GENERAL INFORMATION:
 ; APPLICANT: Lee, James
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: VEGF-RELATED PROTEIN
 ; FILE REFERENCE: P0963R1D1
 ; CURRENT APPLICATION NUMBER: US/09/313,299B
 ; CURRENT FILING DATE: 1999-05-17
 ; EARLIER APPLICATION NUMBER: US 08/706,054
 ; EARLIER FILING DATE: 1996-08-30
 ; EARLIER APPLICATION NUMBER: US 60/003,491
 ; EARLIER FILING DATE: 1995-09-08
 ; NUMBER OF SEQ ID NOS: 12
 ; SEQ ID NO 2
 ; LENGTH: 2031
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: Human
 ; LOCATION: 1-2031
 ; OTHER INFORMATION: Sequence source: complement to SEQ ID NO. 1
 ; Patent No. 6576608
 ; US-09-313-299-2

Query Match 99.3%; Score 1663; DB 4; Length 2031;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1666; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GTCCTTCCACCATGCACTGCTGGCTTCTTCTCTGTGGCGGTGTTCTCTGTCTCCCGCTG 60
 Db 1671 GTCCTTCCACCATGCACTGCTGGCTTCTTCTCTGTGGCGGTGTTCTCTGTCTCCCGCTG 1612
 QY 61 CGCTGCTCCCGGCTCTCTCGCGAGGCGCCGCGCGCGCGCGCGCTTCCGAGTCCGACTCG 120
 Db 1611 TGCTGCTCCCGGCTCTCTCGCGAGGCGCCGCGCGCGCGCGCTTCCGAGTCCGACTCG 1552
 QY 121 ACCTCTCGAGCGCGGAGCGCGCGCGCGCGCGCGCGCTTATGCAAGCAAAAGATCTGG 180
 Db 1551 ACCTCTCGAGCGCGGAGCGCGCGCGCGCGCGCGCGCTTATGCAAGCAAAAGATCTGG 1492
 QY 181 AGGAGCAGTTACGCTGCTGTCCAGTGTAGTGAACATCATGACTGTACTCTTACCCAGAT 240
 Db 1491 AGGAGCAGTTACGCTGCTGTCCAGTGTAGTGAACATCATGACTGTACTCTTACCCAGAT 1432
 QY 241 ATTGGAATGTACAAAGTGTGCTGCTGAGGAAAGAGGCTGCGCAACATAACAGAGAACAGG 300
 Db 1431 ATTGGAATGTACAAAGTGTGCTGCTGAGGAAAGAGGCTGCGCAACATAACAGAGAACAGG 1372
 QY 301 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGCGAGCACATTTATAATACAG 360
 Db 1371 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGCGAGCACATTTATAATACAG 1312
 QY 361 AGATCTTGAAGATTTGATTAATGATGAGAGAAAGACTCAATGATGCTGCGAGGAGTGT 420
 Db 1311 AGATCTTGAAGATTTGATTAATGATGAGAGAAAGACTCAATGATGCTGCGAGGAGTGT 1252
 QY 421 GTATAGATGTGGGGAAGAGATTTGGAGTGGCGCAACACACCTTCTTTAAACCTCCATCTG 480
 Db 1251 GTATAGATGTGGGGAAGAGATTTGGAGTGGCGCAACACACCTTCTTTAAACCTCCATCTG 1192
 QY 481 TGTCCGCTTACAGATGTGGGAGTGTCTGCTCAATAGTGGGAGCTGAGTGAACACCA 540
 Db 1191 TGTCCGCTTACAGATGTGGGAGTGTCTGCTCAATAGTGGGAGCTGAGTGAACACCA 1132
 QY 541 GCACGAGCTTACCTCAGCAAGAGCTTATTGAAATTTACAGTGCCTCTCTCTCAAGGCCCCA 600
 Db 1131 GCACGAGCTTACCTCAGCAAGAGCTTATTGAAATTTACAGTGCCTCTCTCTCAAGGCCCCA 1072
 QY 601 AACCCAGTAACTCAGTTTGGCAATCAGACTTCTGCGGATGATGCTTAAATCTGATG 660
 Db 1071 AACCCAGTAACTCAGTTTGGCAATCAGACTTCTGCGGATGATGCTTAAATCTGATG 1012
 QY 661 TTTACAGACAAGTTCATTTCCATTTATAGACGTTCCCTGCCAGCAACACTACCAAGTGC 720
 Db 1011 TTTACAGACAAGTTCATTTCCATTTATAGACGTTCCCTGCCAGCAACACTACCAAGTGC 952
 QY 721 AGGCGAGNACAGAGCTTCCCGCCACCAATTCATGTTGAATATCAATCATCTGCGAGTGC 780
 Db 951 AGGCGAGNACAGAGCTTCCCGCCACCAATTCATGTTGAATATCAATCATCTGCGAGTGC 892
 QY 781 TGGCTCAGGAAGATTTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGATTC 840
 Db 891 TGGCTCAGGAAGATTTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGATTC 832
 QY 841 ATGACATCTGTGGACCAACAGAGGCTGATGAGAGACTGCTGCTGCTGCTGAGAG 900
 Db 831 ATGACATCTGTGGACCAACAGAGGCTGATGAGAGACTGCTGCTGCTGCTGAGAG 772
 QY 901 CGGGGCTTCCGCTGCGAGCTGTGGACCCCAAGAACTAGACAGAAACTCATGTCAGT 960
 Db 771 CGGGGCTTCCGCTGCGAGCTGTGGACCCCAAGAACTAGACAGAAACTCATGTCAGT 712
 QY 961 GTGTCTGTAAAAAACAACCTCTTCCCGAGCAATGTGGGCGCAACCCAGAGATTTGATGAAA 1020
 Db 711 GTGTCTGTAAAAAACAACCTCTTCCCGAGCAATGTGGGCGCAACCCAGAGATTTGATGAAA 652
 QY 1021 ACACATGCGAGTGTGTATGTAAAAAGAACTGCTCCCGAGAAATCAACCCCTAAATCTTGAA 1080
 Db 651 ACACATGCGAGTGTGTATGTAAAAAGAACTGCTCCCGAGAAATCAACCCCTAAATCTTGAA 592
 QY 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTGTTAAAGGAAAGAGTTCC 1140

591	Db	AAATGTGCGCTGTGAATGTACAGAAAGTCCACAGAAATCTTGTGTTAAAGGAAAGAGTTCC	532
1141	Qy	ACCACCAAAACATGACAGCTGTTACAGACGGCCATGTACAAACCGCCAGAGCGCTTGTGAGC	1200
531	Db	ACCACCAAAACATGACAGCTGTTACAGACGGCCATGTACAAACCGCCAGAGCGCTTGTGAGC	472
1201	Qy	CAGGATTTTCATATAGTGAAGAGTGTGCTGTGTCCTTCATATGTCGCAAGACCAAC	1360
471	Db	CAGATTTTCATATAGTGAAGAGTGTGCTGTGTCCTTCATATGGAAGACCAAC	412
1261	Qy	AAATGAGCTAAGATGTACTGTTTCCAGTTTCATCGATTTTCTATTATGGAACACTGTGT	1320
411	Db	AAATGAGCTAAGATGTACTGTTTCCAGTTTCATCGATTTTCTATTATGGAACACTGTGT	352
1321	Qy	TGCCACAGTAGAAGCTGTCTGTAACACAGAGAGACCCCTTGTGGTCCATGCTAAACAAGACA	1380
351	Db	TGCCACAGTAGAAGCTGTCTGTAACACAGAGAGACCCCTTGTGGTCCATGCTAAACAAGACA	292
1381	Qy	AAAGTCTGTCTTTTCCCTGAAACCATGTGGATAACTTTACAGAAAAATGGACTCGAGCTCATCTG	1440
291	Db	AAAGTCTGTCTTTTCCCTGAAACCATGTGGATAACTTTACAGAAAAATGGACTCATCTG	232
1441	Qy	CAAAAGGCCCTCTGTAAAGACTGGTTTTCGCCAATGACCAAAACAGCCAGATTTTCCCTC	1500
231	Db	CAAAAGGCCCTCTGTAAAGACTGGTTTTCGCCAATGACCAAAACAGCCAGATTTTCCCTC	172
1501	Qy	TTGTGATTTCTTTAAAGAGATGACTATATAATTTATTTTCCACTTAAANAATATTGTTTCTGC	1560
171	Db	TTGTGATTTCTTTAAAGAGATGACTATATAATTTATTTTCCACTTAAANAATATTGTTTCTGC	112
1561	Qy	ATTTCATTTTATAGCAACAACATTTGGTAAAACTCAGTGTATCAATATTTTATATATCAT	1620
111	Db	ATTTCATTTTATAGCAACAACATTTGGTAAAACTCAGTGTATCAATATTTTATATATCAT	52
1621	Qy	GC AAAATATGTTTAAAAATAAAATGAAATTTGTTATTTATAAAAAA AAAAAA 1671	
51	Db	GC AAAATATGTTTAAAAATAAAATGAAATTTGTTATTTAAAAA AAAAAA AAAAAA 1	

RESULT 8

```

PC7-US96-09001-1
:
: Sequence 1, Application PC/TUS9609001
: GENERAL INFORMATION:
: APPLICANT: HU, ET AL.
: TITLE OF INVENTION: Human Vascular En
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: CARELLA, BYRNE, BAIN, GI
: ADDRESSEE: CECCHI, STEWART & OLSTEI
: STREET: 6 BECKER FARM ROAD
: CITY: ROSELAND
: STATE: NEW JERSEY
: COUNTRY: USA
: ZIP: 07068
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 INCH DISKETTE
: OPERATING SYSTEM: MS-DOS
: SOFTWARE: WORD PERFECT 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US96/09001
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/465,968
: FILING DATE: 6 JUN 95
: APPLICATION NUMBER: 08/207,550
: FILING DATE: 8 MAR 1994
: ATTORNEY/AGENT INFORMATION:
: NAME: FERRARO, GREGORY D.
: REGISTRATION NUMBER: 36,134
: REFERENCE/DOCKET NUMBER: 325800-288

```

Db 841 ATGACATCTGTGGACCAAAACAGGAGCTGGATGAAGAGACCTGTGAGTGTGTCTGCAGAG 900
Qy 901 CGGGGCTTGGGCTCGGCTGAGCTGTGGACCCCAAGAGAACTACAGAGAACTCATGCCAGT 960
Db 901 CGGGGCTTGGGCTCGGCTGAGCTGTGGACCCCAAGAGAACTACAGAGAACTCATGCCAGT 960
Qy 961 GTGTCTGTAAAAACAACTCTTCCCGAGCAATGTGGGCGCCCAACCGAGAAATTTGATGAAA 1020
Db 961 GTGTCTGTAAAAACAACTCTTCCCGAGCAATGTGGGCGCCCAACCGAGCAATTTGATGAAA 1020
Qy 1021 ACACATGCCAGTGTGTATGTAAAGAACTGCCCGAGAAATCAACCCCTAAATCTCGAA 1080
Db 1021 ACACATGCCAGTGTGTATGTAAAGAACTGCCCGAGAAATCAACCCCTAAATCTCGAA 1080
Qy 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTTAAAGGAAAGAAAGTTCC 1140
Db 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTTAAAGGAAAGAAAGTTCC 1140
Qy 1141 ACCACCAACATGCAGCTGTACAGAGCGGCATGTACGACCGCCAGAGGCTTGTGAGC 1200
Db 1141 ACCACCAACATGCAGCTGTACAGAGCGGCATGTACGACCGCCAGAGGCTTGTGAGC 1200
Qy 1201 CAGGATTTTCATATAGTGAAGAGTGTGTGCTGCTCCCTTCATATGCGCAAGACCAAC 1260
Db 1201 CAGGATTTTCATATAGTGAAGAGTGTGTGCTGCTCCCTTCATATGCGCAAGACCAAC 1260
Qy 1261 AATGAGCTAAGATGTATGCTGTTTCCAGTTCATCGATTTTCTATATGGAAGAACTGTGT 1320
Db 1261 AATGAGCTAAGATGTATGCTGTTTCCAGTTCATCGATTTTCTATATGGAAGAACTGTGT 1320
Qy 1321 TGCCACAGTAGAAGTGTGTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 1321 TGCCACAGTAGAAGTGTGTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Qy 1381 AAAGTCTGCTTTTCTGAACCAATGTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 1381 AAAGTCTGCTTTTCTGAACCAATGTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Qy 1441 CAAAGGCTTCTGTAAG 1500
Db 1441 CAAAGGCTTCTGTAAG 1500
Qy 1501 TTGTGATTTCTTTAAAG 1560
Db 1501 TTGTGATTTCTTTAAAG 1560
Qy 1561 ATTCAATTTTATAGCAACAAATGTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
Db 1561 ATTCAATTTTATAGCAACAAATGTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
Qy 1621 GCAAAATATCTTTAAAG 1674
Db 1621 GCAAAATATCTTTAAAG 1674

RESULT 9
US-08-795-430-7
; Sequence 7, Application US/08795430
; Patent No. 6130071
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Joukov, Vladimir
; TITLE OF INVENTION: Vascular Endothelial Growth Factor C (VEGF-C)
; TITLE OF INVENTION: Protein and Gene, Mutants Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/795,430
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FI96/00427
FILING DATE: 01-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/671,573
FILING DATE: 28-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/601,132
FILING DATE: 14-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/585,895
FILING DATE: 12-JAN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/510,133
FILING DATE: 01-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,011
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28967/33691
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1997 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 352..1608
US-08-795-430-7

Query Match 98.7%; Score 1651.8; DB 3; Length 1997;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 GTCCTTCCACCATGCACTCGCTGGGCTTCTTCTGTGGGCTGTCTCTCTCGCCGCTG 60
Db 341 GTCCTTCCACCATGCACTTGTGGGCTTCTTCTGTGGGCTGTCTCTCTCGCCGCTG 400
Qy 61 CGCTGCTCCCGGCTCCTCGGAGGCGCCGCGCGCGCGCGCTTCGAGTCCGAGCTCG 120
Db 401 CGCTGCTCCCGGCTCCTCGGAGGCGCCGCGCGCGCGCTTCGAGTCCGAGCTCG 460
Qy 121 ACCTCTCGGAGCGCGAGCGCCGAGCGCGCGCGCGCTTCGAGTCCGAGCTCG 180
Db 461 ACCTCTCGGAGCGCGAGCGCCGAGCGCGCGCGCTTCGAGTCCGAGCTCG 520
Qy 181 AGGAGCGAGTTCAGGCTGTGTGTCCAGTGTAGATGAGTCACTGAGTGTACTTACCCAGAT 240
Db 521 AGGAGCGAGTTCAGGCTGTGTGTCCAGTGTAGATGAGTCACTGAGTGTACTTACCCAGAT 580
Qy 241 ATTGGAATGTACAAAGTGTGAGTAAAGAGAGAGCTGCGCAACATAACAGAGAACAGG 300
Db 581 ATTGGAATGTACAAAGTGTGAGTAAAGAGAGAGCTGCGCAACATAACAGAGAACAGG 640
Qy 301 CCAACCTCACTCAAGGACAGAGAGAGATATAAATTTGCTGAGGACATTTATTAATACAG 360

Db 641 CCAACCTCAACTCAAGGACAGAGAGACTATAAAATTTGCTGCGACACATTAATAACAG 700
 Qy 361 AGATCTTTGAAAGTATTGATTAATCAGTGGAGAAAGACTCAATGATGCCACCGGAGGTGT 420
 Db 701 AGATCTTTGAAAGTATTGATTAATCAGTGGAGAAAGACTCAATGATGCCACCGGAGGTGT 760
 Qy 421 GTATAGATGTGGGAAAGAGTGTGGAGTCGGGCAAAACACCTTTCTTTAAACCTCCATGTG 480
 Db 761 GTATAGATGTGGGAAAGAGTGTGGAGTCGGGCAAAACACCTTTCTTTAAACCTCCATGTG 820
 Qy 481 TGTCCGCTCTACAGATGTGGGCTTGTCTGCAATAGTAGGGGCTGCAATGATGAACACCA 540
 Db 821 TGTCCGCTCTACAGATGTGGGCTTGTCTGCAATAGTAGGGGCTGCAATGATGAACACCA 880
 Qy 541 GCACGAGTACCTCAGCAAGAGCTTATTTGAATTAAGTGCCTCTCTCAAGGCCCCA 600
 Db 881 GCACGAGTACCTCAGCAAGAGCTTATTTGAATTAAGTGCCTCTCTCAAGGCCCCA 940
 Qy 601 AACGAGTAACATCAGTTTTCGCAATCACATCTCCGATGATGCTTAAACTGATG 660
 Db 941 AACGAGTAACATCAGTTTTCGCAATCACATCTCCGATGATGCTTAAACTGATG 1000
 Qy 661 TTTACAGCAAGTTCATTCCATTATTAGAGCTTCCCTGCCAGCAACACTACCAAGTGC 720
 Db 1001 TTTACAGCAAGTTCATTCCATTATTAGAGCTTCCCTGCCAGCAACACTACCAAGTGC 1060
 Qy 721 AGGAGGAAACAGACCTGCCACCAATACATGTGGATTAATCACATCTGCAGATGCC 780
 Db 1061 AGGAGGAAACAGACCTGCCACCAATACATGTGGATTAATCACATCTGCAGATGCC 1120
 Qy 781 TGGCTCAGAAAGATTTATGTTTTCTCGATGCTGGAGATGACTCAACAGATGGATTCC 840
 Db 1121 TGGCTCAGAAAGATTTATGTTTTCTCGATGCTGGAGATGACTCAACAGATGGATTCC 1180
 Qy 841 ATGACATCTGAGACCAACAAAGAGCTGGATGAAGAGACCTGTCACTGTCTGCGAG 900
 Db 1181 ATGACATCTGAGACCAACAAAGAGCTGGATGAAGAGACCTGTCACTGTCTGCGAG 1240
 Qy 901 CGGGCTTCGGCTCGCAGCTGTGACCCCAAGAACTAGACAGAACTCATGCCAGT 960
 Db 1241 CGGGCTTCGGCTCGCAGCTGTGACCCCAAGAACTAGACAGAACTCATGCCAGT 1300
 Qy 961 GTGCTGTAAACAAACTCTTCCCAAGCAATGTGGGCCAACCGAGAAATTTGATGAAA 1020
 Db 1301 GTGCTGTAAACAAACTCTTCCCAAGCAATGTGGGCCAACCGAGAAATTTGATGAAA 1360
 Qy 1021 ACACATGCGAGTGTATGTAAGAACTGCCAGCAAACTCAACCCCTAACTCCTGGA 1080
 Db 1361 ACACATGCGAGTGTATGTAAGAACTGCCAGCAAACTCAACCCCTAACTCCTGGA 1420
 Qy 1081 AATGTCCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTTAAAGGAAAGAGTTCC 1140
 Db 1421 AATGTCCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTTAAAGGAAAGAGTTCC 1480
 Qy 1141 ACCACCAACATGAGCTGTTACAGAGGCAATGACGAGCCGACAGAGCTTTGAGC 1200
 Db 1481 ACCACCAACATGAGCTGTTACAGAGGCAATGACGAGCCGACAGAGCTTTGAGC 1540
 Qy 1201 CAGGATTTTCATATAGTGAAGAGTGTCTGTTGTCCTTTCATATTTGGCAAGACAC 1260
 Db 1541 CAGGATTTTCATATAGTGAAGAGTGTCTGTTGTCCTTTCATATTTGGCAAGACAC 1600
 Qy 1261 AAATGAGCTAAGATTTACTGTTTTCCAGTTTCATGATTTTCTATTTATGGAATCTGTGT 1320
 Db 1601 AAATGAGCTAAGATTTACTGTTTTCCAGTTTCATGATTTTCTATTTATGGAATCTGTGT 1660
 Qy 1321 TGCCACAGTGAAGTGTCTGTGAACAGAGAGCCCTTGTGGGTCCATGCTAACAAAGACA 1380
 Db 1661 TGCCACAGTGAAGTGTCTGTGAACAGAGAGCCCTTGTGGGTCCATGCTAACAAAGACA 1720
 Qy 1381 AAAGTCTGTCTTCTGAAACCAATGTGGATAACTTTTACAGAAATGAGCTTGGAGCTCATCTG 1440
 Db 1721 AAAGTCTGTCTTCTGAAACCAATGTGGATAACTTTTACAGAAATGAGCTTGGAGCTCATCTG 1780

Qy 1441 CAAAAGGCTCTTTGAAAGACTGGTTTTCTGCCAATGACCAAAACAGCAAGATTTTCTC 1500
 Db 1781 CAAAAGGCTCTTTGAAAGACTGGTTTTCTGCCAATGACCAAAACAGCAAGATTTTCTC 1840
 Qy 1501 TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAAAATATTGTTCTGC 1560
 Db 1841 TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAAAATATTGTTCTGC 1900
 Qy 1561 ATTCAATTTTATAGCAACAAATTTGGTAAACTCCTGTGATCAATATTTTATATCAT 1620
 Db 1901 ATTCAATTTTATAGCAACAAATTTGGTAAACTCCTGTGATCAATATTTTATATCAT 1960
 Qy 1621 GCAAAATGTTTAAATTAATGAAATTTGTTT 1655
 Db 1961 GCAAAATGTTTAAATTAATGAAATTTGTTT 1995

RESULT 10

US-08-510-133A-34
 ; Sequence 34, Application US/08510133A
 ; Patent No. 6221839
 ; GENERAL INFORMATION:
 ; APPLICANT: Alitalo, Kari
 ; Jukov, Vladimir
 ; TITLE OF INVENTION: Receptor Ligand
 ; NUMBER OF SEQUENCES: 35
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/510,133A
 ; FILING DATE: 01-Aug-1995
 ; CLASSIFICATION: <unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Gass, David A.
 ; REGISTRATION NUMBER: 38,153
 ; REFERENCE/DOCKET NUMBER: 28113/32863
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312/474-6300
 ; TELEFAX: 312/474-0446
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 34:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1997 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 352..1608
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 34:
 US-08-510-133A-34

Query Match 98.7%; Score 1651.8; DB 3; Length 1997;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 GTCTTCCACCATGCACTCGCTGGGCTTCTTCTGTGGCGGTCTTCTGTGCTCGCGCTG 60
 Db 341 GTCTTCCACCATGCACTGCTGGGCTTCTTCTGTGGCGGTCTTCTGTGCTCGCGCTG 400
 Qy 61 CGCTGTCTCCGGGTCTCTCGAGAGCGCCCGCCCGCTTTCGAGTCGGAGACTCG 120

Db 401 CGCTGCTCCGGGTCCTTCGAGGGCCCGCCGCGCGCTTCGAGTCGGACTCG 450
Qy 121 ACCTCTCGGACGGGAGCCCGAGCGCGGCGAGCCACGGCTTATGCAAGCAAGATCTGG 180
Db 461 ACCTCTCGGACGGGAGCCCGAGCGCGGCGAGCCACGGCTTATGCAAGCAAGATCTGG 520
Qy 181 AGGAGCAGTTACGGTCTGTGTCAGTGTAGATGAATCTCATGCTGTACTCTACCCAGAT 240
Db 521 AGGAGCAGTTACGGTCTGTGTCAGTGTAGATGAATCTCATGCTGTACTCTACCCAGAT 580
Qy 241 ATTGGAAAAATGTACAAGTGTGAGTAAAGGAGGCTGGCAACATAACAGAGAAACAG 300
Db 581 ATTGGAAAAATGTACAAGTGTGAGTAAAGGAGGCTGGCAACATAACAGAGAAACAG 640
Qy 301 CCAACCTCAACTCAAGACAGAGAGACTATAAAATTGCTGAGCAGCAGCATTAATAACAG 360
Db 641 CCAACCTCAACTCAAGACAGAGAGACTATAAAATTGCTGAGCAGCAGCATTAATAACAG 700
Qy 361 AGATCTTTGAAAAATGATAAATGAGTGGAGAAAGACTCAATGATGCCACCGGAGGTGT 420
Db 701 AGATCTTTGAAAAATGATAAATGAGTGGAGAAAGACTCAATGATGCCACCGGAGGTGT 760
Qy 421 GTATAGATGTGGGAGAGAGTTGGAGTCCGGAACAAACACTTCTTTAAACCTCCATGTG 480
Db 761 GTATAGATGTGGGAGAGAGTTGGAGTCCGGAACAAACACTTCTTTAAACCTCCATGTG 820
Qy 481 TGTCCGCTACAGATGTGGGGTGTGCTCAATAGTGGGGGCTGCAGTGCATGAACACCA 540
Db 821 TGTCCGCTACAGATGTGGGGTGTGCTCAATAGTGGGGGCTGCAGTGCATGAACACCA 880
Qy 541 GCACGAGTACCTCAGCAAGACGTTATTTGAAATTAACAGTGCCTCTCTCTCAAGGCCCA 600
Db 881 GCACGAGTACCTCAGCAAGACGTTATTTGAAATTAACAGTGCCTCTCTCTCAAGGCCCA 940
Qy 601 AACAGTAAACATCAGTTTGGCAATCACACTTCTCCGATGCATGCTCTAACTCGATG 660
Db 941 AACAGTAAACATCAGTTTGGCAATCACACTTCTCCGATGCATGCTCTAACTCGATG 1000
Qy 561 TTTACAGACAGTTTCATTTCATTTAGAGCTTCCCTGCCAGCAACACTACCAAGTGTG 720
Db 1001 TTTACAGACAGTTTCATTTCATTTAGAGCTTCCCTGCCAGCAACACTACCAAGTGTG 1060
Qy 721 AGGAGCGAACAAGACTGCCCCCAATACATGTGGATATCATCTGCAGATGCC 780
Db 1061 AGGAGCGAACAAGACTGCCCCCAATACATGTGGATATCATCTGCAGATGCC 1120
Qy 781 TGGCTCAGGAAGATTTATGTTTCTCGATGCTGGAGATGACTCAACAGATGGATTC 840
Db 1121 TGGCTCAGGAAGATTTATGTTTCTCGATGCTGGAGATGACTCAACAGATGGATTC 1180
Qy 841 ATGACATCTGTGGACCAAAACAGAGCTGGATGAAGAGACTGTGATGCTGTCTGCAGAG 900
Db 1181 ATGACATCTGTGGACCAAAACAGAGCTGGATGAAGAGACTGTGATGCTGTCTGCAGAG 1240
Qy 901 CGGSGCTTCGCGCTCGCAGCTGTGAGCCCAAGAACTAGACAGAAACTCATGCCAGT 960
Db 1241 CGGSGCTTCGCGCTCGCAGCTGTGAGCCCAAGAACTAGACAGAAACTCATGCCAGT 1300
Qy 961 GTGTCTGTAAAAAACAACCTCTCCCGAGCAATGTGGGCAACCGAGAAATTTGATGAAA 1020
Db 1301 GTGTCTGTAAAAAACAACCTCTCCCGAGCAATGTGGGCAACCGAGAAATTTGATGAAA 1360
Qy 1021 ACACATGCGAGTGTATGTATTAAGAAACCTGCCCCAGAAATCAACCCCTAAATCTCGAA 1080
Db 1361 ACACATGCGAGTGTATGTATTAAGAAACCTGCCCCAGAAATCAACCCCTAAATCTCGAA 1420
Qy 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTGTTAAAAAGGAAAGAGTTC 1140
Db 1421 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTGTTAAAAAGGAAAGAGTTC 1480
Qy 1141 ACCACCAACATGCGAGTGTACAGACGGCCATGTACGAACCGCCAGAGGCTTGTGAGC 1200

Db 1481 ACCACCAACATGCGAGTGTACAGACGGCCATGTACGAACCGCCAGAGGCTTGTGAGC 1540
Qy 1201 CAGAAATTTTCATATAGTGAAGAGTGTGTCTGTTGTCTCCCTTCATATTTGGCAAGACAC 1260
Db 1541 CAGAAATTTTCATATAGTGAAGAGTGTGTCTGTTGTCTCCCTTCATATTTGGAAAGACAC 1600
Qy 1261 AAATGAGCTAAGATTTACTGTTTCCAGTTTCATCGATTTTCTATTTATGGAAAACTGTG 1320
Db 1601 AAATGAGCTAAGATTTACTGTTTCCAGTTTCATCGATTTTCTATTTATGGAAAACTGTG 1660
Qy 1321 TGCCACAGTGAACCTGCTGTGAACAGAGAGACCTTGTGGGTCCATGCTTAACAAAGACA 1380
Db 1661 TGCCACAGTGAACCTGCTGTGAACAGAGAGACCTTGTGGGTCCATGCTTAACAAAGACA 1720
Qy 1381 AAAGTCTCTCTTCTGAAACCATGTGGATAACTTTACAGAAATGGAGTGGAGCTCATCTG 1440
Db 1721 AAAGTCTCTCTTCTGAAACCATGTGGATAACTTTACAGAAATGGAGTGGAGCTCATCTG 1780
Qy 1441 CAAAGGCTCTTGTAAAGACTGTTTCTGCCAATGACCAACAGCAAGATTTTCTCTC 1500
Db 1781 CAAAGGCTCTTGTAAAGACTGTTTCTGCCAATGACCAACAGCAAGATTTTCTCTC 1840
Qy 1501 TTGTGATTTCTTTAAAGAACTGATATATATTTTCCACTAAATAATTTGTTCTGCG 1560
Db 1841 TTGTGATTTCTTTAAAGAACTGATATATATTTTCCACTAAATAATTTGTTCTGCG 1900
Qy 1561 ATTCAATTTTATAGCAACACAAATTTGGTAAACACTCAGTGTGATCAATATTTTATATCAT 1620
Db 1901 ATTCAATTTTATAGCAACACAAATTTGGTAAACACTCAGTGTGATCAATATTTTATATCAT 1960
Qy 1621 GCAAAATATGTTTAAATAAATGAATGAAATTTGTAAT 1655
Db 1961 GCAAAATATGTTTAAATAAATGAATGAAATTTGTAAT 1995

RESULT 11
US-09-355-700-7
; Sequence 7, Application US/09355700
; Patent No. 6361946
; GENERAL INFORMATION:
; APPLICANT: Ludwag Institute for Cancer Research
; Heisinki University Licensing
; Alitalo, Kari (U.S. only)
; Joukov, Vladimir (U.S. only)
; TITLE OF INVENTION: Vascular Endothelial Growth Factor C (VEGF-C)
; Protein and Gene, Mutants Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/355,700
; FILING DATE: 05-NOV-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/795,430
; FILING DATE: 05-FEB-1997
; APPLICATION NUMBER: PCT/PT96/00427
; FILING DATE: 01-AUG-1996
; APPLICATION NUMBER: 08/671,573
; FILING DATE: 28-JUN-1996
; APPLICATION NUMBER: 08/601,132
; FILING DATE: 14-FEB-1996
; APPLICATION NUMBER: 08/585,895

```

FILING DATE: 12-JAN-1996
APPLICATION NUMBER: 08/510,133
FILING DATE: 01-AUG-1995
APPLICATION NUMBER: 08/340,011
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28967/34140
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1997 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: CDS
LOCATION: 352..1608
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-355-700-7

Query Match
Best Local Similarity 99.7%; Score 1651.8; DB 4; Length 1997;
Matches 1653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTCCCTCCACATGCACTCTGGGCTTCTCTGCTGGGCTGTTCTCTGCTCGCGCTG 60
DB 341 GTCCCTCCACATGCACTCTGGGCTTCTCTGCTGGGCTGTTCTCTGCTCGCGCTG 400
QY 61 CGCTGCTCCGCGGTCCTCGGAGCGCCGCGCGCGCGCGCGCTTCGAGTCCGAGCTCG 120
DB 401 CGCTGCTCCGCGGTCCTCGGAGCGCCGCGCGCGCGCGCTTCGAGTCCGAGCTCG 460
QY 121 ACCTCTCGAGCGCGAGCGCGAGCGCGCGCGCGCGCGCTTCGAGTCCGAGCTCG 180
DB 461 ACCTCTCGAGCGCGAGCGCGAGCGCGCGCGCGCGCGCTTCGAGTCCGAGCTCG 520
QY 181 AGGAGCAGTTACGCTGTGTGTCAGTGTAGATGAATCATGACTGACTCTACCCAGAT 240
DB 521 AGGAGCAGTTACGCTGTGTGTCAGTGTAGATGAATCATGACTGACTCTACCCAGAT 580
QY 241 ATTGGAAATGTAACAAGTGTGCTGTCAGTGTAGATGAATCATGACTGACTCTACCCAGAT 300
DB 581 ATTGGAAATGTAACAAGTGTGCTGTCAGTGTAGATGAATCATGACTGACTCTACCCAGAT 640
QY 301 CCACCTCACTCAGGACGAGAGACTATAAATTTGCTGAGACACATTATAATACAG 360
DB 641 CCACCTCACTCAGGACGAGAGACTATAAATTTGCTGAGACACATTATAATACAG 700
QY 361 AGATCTTGAAGAAGTATTGATTAATGAGTGGAGAAAGACTCAATGATCCCGAGAGTGT 420
DB 701 AGATCTTGAAGAAGTATTGATTAATGAGTGGAGAAAGACTCAATGATCCCGAGAGTGT 760
QY 421 GTATAGATGTTGGGAAGAGTTGGAGTTCGCGCAACACCTTCTTAAACCTCCATGTG 480
DB 761 GTATAGATGTTGGGAAGAGTTGGAGTTCGCGCAACACCTTCTTAAACCTCCATGTG 820
QY 481 TGTCCGTCTACAGATGTGGGGGTTGCTGCAATAGTGAAGGGGCTGAGTGCATGAACACCA 540
DB 821 TGTCCGTCTACAGATGTGGGGGTTGCTGCAATAGTGAAGGGGCTGAGTGCATGAACACCA 880
QY 541 GCACGAGTCACTCAGCAAGAGTATTGTAATACAGTGCCTCTCTCAGGCGCCCA 600
DB 881 GCACGAGTCACTCAGCAAGAGTATTGTAATACAGTGCCTCTCTCAGGCGCCCA 940
QY 601 AACAGTAAATCAATGTTTTCGCAATCACTCTCTGCGGATGATGCTCAAACTGGATG 660
DB 941 AACAGTAAATCAATGTTTTCGCAATCACTCTCTGCGGATGATGCTCAAACTGGATG 1000

```

RESULT 12
US-08-601-132-32
; Sequence 32, Application US/08601132

```

QY 661 TTTACAGACAAAGTTTCATTTCATTATTAGACGTTCCCTGCCAGCAACTACCAAGTGTG 720
DB 1001 TTTACAGACAAAGTTTCATTTCATTATTAGACGTTCCCTGCCAGCAACTACCAAGTGTG 1060
QY 721 AGGACAGCAAGACAGACCTGCCACCAATACATGTGGTAATACATCTGCGAGATGCC 780
DB 1061 AGGACAGCAAGACAGACCTGCCACCAATACATGTGGTAATACATCTGCGAGATGCC 1120
QY 781 TGGCTCAGGAAGATTATTTATGTTTTCCTCGGATGCTGGAGATGACTCAACAGATGGATTC 840
DB 1121 TGGCTCAGGAAGATTATTTATGTTTTCCTCGGATGCTGGAGATGACTCAACAGATGGATTC 1180
QY 841 ATGACATCTGTGGACCAAAACAAGAGCTGGATGTAAGAGACCTGTCAGTGTCTGCGAGAG 900
DB 1181 ATGACATCTGTGGACCAAAACAAGAGCTGGATGTAAGAGACCTGTCAGTGTCTGCGAGAG 1240
QY 901 CGGGCTTCGGCTGCCAGCTGTGGACCCCAAAAGAACTAGACAGAGAACTCATGCCAGT 960
DB 1241 CGGGCTTCGGCTGCCAGCTGTGGACCCCAAAAGAACTAGACAGAGAACTCATGCCAGT 1300
QY 961 GTGCTGTAAAAACAAACTCTTCCCGAGCCAAATGTGGGGCCCAACCGAGAAATTTGATGAA 1020
DB 1301 GTGCTGTAAAAACAAACTCTTCCCGAGCCAAATGTGGGGCCCAACCGAGAAATTTGATGAA 1360
QY 1021 ACACATGCCAGTGTGTATGTAAGAAAGAACTGCCCGAGAAATCAACCCCTAAATCTCGAA 1080
DB 1361 ACACATGCCAGTGTGTATGTAAGAAAGAACTGCCCGAGAAATCAACCCCTAAATCTCGAA 1420
QY 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTTAAAGGAAAGAAAGTTCC 1140
DB 1421 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTTAAAGGAAAGAAAGTTCC 1480
QY 1141 ACCACAAACATGAGCTGTTTACAGAGCGCCATGAGAACCGCCAGAGAGCTGTGAGC 1200
DB 1481 ACCACAAACATGAGCTGTTTACAGAGCGCCATGAGAACCGCCAGAGAGCTGTGAGC 1540
QY 1201 CAGGATTTTCATATAGTGAAGAGTGTGTCGTTGTGTCCTTCATATTTGGCAAGACAC 1260
DB 1541 CAGGATTTTCATATAGTGAAGAGTGTGTCGTTGTGTCCTTCATATTTGGCAAGACAC 1600
QY 1261 AATGAGCTAGATTTGTAAGTCTGTTTCCAGTTCATCGATTTTCTATTATGGAAGAACTGTG 1320
DB 1601 AATGAGCTAGATTTGTAAGTCTGTTTCCAGTTCATCGATTTTCTATTATGGAAGAACTGTG 1660
QY 1321 TGCCACAGTGAAGTCTGTTGAAACAGAGAGACCTGTCGCTCCATGCTTAACAAAGACA 1380
DB 1661 TGCCACAGTGAAGTCTGTTGAAACAGAGAGACCTGTCGCTCCATGCTTAACAAAGACA 1720
QY 1381 AAAGTCTGCTTTTCTGAAACCATGTTGATACTTTTACAGAAATGAGCTGGAGCTCATCTG 1440
DB 1721 AAAGTCTGCTTTTCTGAAACCATGTTGATACTTTTACAGAAATGAGCTGGAGCTCATCTG 1780
QY 1441 CAAAAGGCTCTGTTGAAAGACTGTTTTCGCAATGACCAACAGCAAGATTTTCTCTC 1500
DB 1781 CAAAAGGCTCTGTTGAAAGACTGTTTTCGCAATGACCAACAGCAAGATTTTCTCTC 1840
QY 1501 TTGTGATTTCTTTAAAGAAAGTACTATATAATTTATTTCCACTAAAAATAATGTTTCTGC 1560
DB 1841 TTGTGATTTCTTTAAAGAAAGTACTATATAATTTATTTCCACTAAAAATAATGTTTCTGC 1900
QY 1561 ATTCAATTTATAGCAACAAAGTGTGTAAGAACTCACTGTGATCAATTTTTTATATCAT 1620
DB 1901 ATTCAATTTATAGCAACAAAGTGTGTAAGAACTCACTGTGATCAATTTTTTATATCAT 1660
QY 1621 GCAAAATATGTTTAAAAATAAAATGAAATTTGATT 1655
DB 1961 GCAAAATATGTTTAAAAATAAAATGAAATTTGATT 1995

```


Db 1781 CAAAAGGCTCTGTGAAAGATGGTTTTCTGCCAATGACCAACAGCCAGATTTCTCTC 1840
Qy 1501 TTGTGATTTCTTTAAAGAAAGACTATATAATTTATTTCCACTAAATAATTTGTTCTGC 1560
Db 1841 TTGTGATTTCTTTAAAGAAAGACTATATAATTTATTTCCACTAAATAATTTGTTCTGC 1900
Qy 1561 ATTCAATTTTATAGCAACAAATGTTGTAAGAACTCAGTGATCAATATTTTATATCAT 1620
Db 1901 ATTCAATTTTATAGCAACAAATGTTGTAAGAACTCAGTGATCAATATTTTATATCAT 1960
Qy 1621 GCAAAATATGTTTAAATAAATAAATGAATGAAATTTGATT 1655
Db 1961 GCAAAATATGTTTAAATAAATAAATGAATGAAATTTGATT 1995

RESULT 13
US-08-671-573B-32
; Sequence 32, Application US/08671573B
; Patent No 6645933
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Joukov, Vladimir
; TITLE OF INVENTION: Receptor Ligand
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/671,573B
; FILING DATE: 28-JUN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/601,132
; FILING DATE: 14-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/585,895
; FILING DATE: 12-JAN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/510,133
; FILING DATE: 01-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Gass, David A.
; REGISTRATION NUMBER: 38,153
; REFERENCE/DOCKET NUMBER: 28967/33348
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1997 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 352..1608
US-08-671-573B-32

Query Match 98.7%; Score 1651.8; DB 4; Length 1997;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GTCTTCCACCATGCACTCGCTGGGCTTCTTCTCTGTGGCGTGTCTCTGTCTCGCGCTG 60
Db 341 GTCTTCCACCATGCACTTGTGGGCTTCTTCTCTGTGGCGTGTCTCTGTCTCGCGCTG 400
Qy 61 CGCTGCTCCGGGTCTCTCGGAGGGCCCGCGCGCGCGCGCGCTTCGAGTTCGGACTCG 120
Db 401 CGCTGCTCCGGGTCTCTCGGAGGGCCCGCGCGCGCGCGCTTCGAGTTCGGACTCG 460
Qy 121 ACCTCTCGGACGCGGAGCCCGGCGGCGGAGGCGGCGGCTTATGCAAGCAAGATCTCG 180
Db 461 ACCTCTCGGACGCGGAGCCCGGCGGCGGAGGCGGCGGCTTATGCAAGCAAGATCTCG 520
Qy 181 AGGAGCAGTTACGCTCTGTGTCTGAGTGTAGATGAATCTCATGACTGTACTCTTACCCAGAT 240
Db 521 AGGAGCAGTTACGCTCTGTGTCTGAGTGTAGATGAATCTCATGACTGTACTCTTACCCAGAT 580
Qy 241 ATTGGAAAATGTACAGTGTCAAGTGAAGGAAAGGAGGCTGGCAACATATAACAGAACAG 300
Db 581 ATTGGAAAATGTACAGTGTCAAGTGAAGGAAAGGAGGCTGGCAACATATAACAGAACAG 640
Qy 301 CCAACCTCAACTCAAGGACAGAGAGACTATATAATTTGCTCGACACATATAATACAG 360
Db 641 CCAACCTCAACTCAAGGACAGAGAGACTATATAATTTGCTCGACACATATAATACAG 700
Qy 361 AGATCTTGAAGATATTGATATGAGTGGAGAAAGACTCAATGCATGCCACGGAGGTCT 420
Db 701 AGATCTTGAAGATATTGATATGAGTGGAGAAAGACTCAATGCATGCCACGGAGGTCT 760
Qy 421 GTATAGATGTGGGAAGGAGTTTGGAGTGCACGACAAAACACCTTCTTTAAACCTCCATGTG 480
Db 761 GTATAGATGTGGGAAGGAGTTTGGAGTGCACGACAAAACACCTTCTTTAAACCTCCATGTG 820
Qy 481 TGTCCGCTTACAGATGTGGGGTTGCTGCAATAGTAGAGGGGCTGCAGTGCATGAACACCA 540
Db 821 TGTCCGCTTACAGATGTGGGGTTGCTGCAATAGTAGAGGGGCTGCAGTGCATGAACACCA 880
Qy 541 GCACGAGCTACCTCAGCAAGAGCTTATTGAAATTTACAGTGCCTCTCTCTCAAGGCCCA 600
Db 881 GCACGAGCTACCTCAGCAAGAGCTTATTGAAATTTACAGTGCCTCTCTCTCAAGGCCCA 940
Qy 601 AACAGTAACAATCAGTTTGGCAATCAACATTCCTGCGGATGCATGTCTAACTGGATG 660
Db 941 AACAGTAACAATCAGTTTGGCAATCAACATTCCTGCGGATGCATGTCTAACTGGATG 1000
Qy 661 TTTACAGACAAAGTTCATTCCTATTAGACGTTCCCTGCGCAGCAACACACTACCACAGTGC 720
Db 1001 TTTACAGACAAAGTTCATTCCTATTAGACGTTCCCTGCGCAGCAACACACTACCACAGTGC 1060
Qy 721 AGCAGCGAACAGACCTCGCCCAATTTACATGTGGAATTAATCACATCTCCAGATGCC 780
Db 1061 AGCAGCGAACAGACCTCGCCCAATTTACATGTGGAATTAATCACATCTCCAGATGCC 1120
Qy 781 TGGCTCAGGAAGATTTTATGTTTCTCGGATGCTGGAGTCACTCAACAGATGATTC 840
Db 1121 TGGCTCAGGAAGATTTTATGTTTCTCGGATGCTGGAGTCACTCAACAGATGATTC 1180
Qy 841 ATGACATCTGTGSAACCAAAAGAGCTGGATGAAGAGACCTGTGAGTGTCTCGCAGAG 900
Db 1181 ATGACATCTGTGSAACCAAAAGAGCTGGATGAAGAGACCTGTGAGTGTCTCGCAGAG 1240
Qy 901 CGGGGCTTCGGCTGCGAGCTGTGGACCCCAAAAGAACTAGACAGAAACTCATGCCAGT 960
Db 1241 CGGGGCTTCGGCTGCGAGCTGTGGACCCCAAAAGAACTAGACAGAAACTCATGCCAGT 1300
Qy 961 GTGCTGTAAAAACAAACTCTTTCCCGAGCAATGTGGGCGCCCAACCGAGAAATTTGATGAA 1020
Db 1301 GTGCTGTAAAAACAAACTCTTTCCCGAGCAATGTGGGCGCCCAACCGAGAAATTTGATGAA 1360
Qy 1021 ACATATGCCAGTGTGTATATAAGAACTGCGCCAGAAATCAACCCCTAAATTCCTGGAA 1080
Db 1361 ACATATGCCAGTGTGTATATAAGAACTGCGCCAGAAATCAACCCCTAAATTCCTGGAA 1420

QY 1081 AATGTCCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAGAGTTCC 1140
Db 1421 AATGTCCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGGAAAGAGTTCC 1480
QY 1141 ACCACCAAAATCAGCTGTACAGACGGCCATGTACGAACCGCCAGAGGCTTGTGAGC 1200
Db 1481 ACCACCAAAATCAGCTGTACAGACGGCCATGTACGAACCGCCAGAGGCTTGTGAGC 1540
QY 1201 CAGGATTTTCATATAGTAAGAGTGTGCTGTGCTCCCTTCATATGCGCAAGACCAAC 1260
Db 1541 CAGGATTTTCATATAGTAAGAGTGTGCTGTGCTCCCTTCATATGCGCAAGACCAAC 1600
QY 1261 AATGAGCTAAGATTTACTGTTTCCAGTTCATGATTTTCTATATGCGCAAGAGTGT 1320
Db 1601 AATGAGCTAAGATTTACTGTTTCCAGTTCATGATTTTCTATATGCGCAAGAGTGT 1660
QY 1321 TGCCACAGTAGAACTGTGTGAACAGAGAGACCCCTTGTGGTCCATGTATACAAAGACA 1380
Db 1661 TGCCACAGTAGAACTGTGTGAACAGAGAGACCCCTTGTGGTCCATGTATACAAAGACA 1720
QY 1381 AAGTCTGTCTTCCCTGAACCATGTGATAACTTTACAGAAATGAGCTGAGCTCATCTG 1440
Db 1721 AAGTCTGTCTTCCCTGAACCATGTGATAACTTTACAGAAATGAGCTGAGCTCATCTG 1780
QY 1441 CAAAGGCTCTTGTAAAGAGTGTGTTTCTGCAATGACCAACAGCCCAAGATTTTCTC 1500
Db 1781 CAAAGGCTCTTGTAAAGAGTGTGTTTCTGCAATGACCAACAGCCCAAGATTTTCTC 1840
QY 1501 TTGTGATTTCTTTAAAGAGTACTATATAATTTATTTCCACTAAATAATTTGTTCTGC 1560
Db 1841 TTGTGATTTCTTTAAAGAGTACTATATAATTTATTTCCACTAAATAATTTGTTCTGC 1900
QY 1561 ATTCAATTTTATAGCAACCAATTTGTTAAACTCACTGTGATCAATTTTATATCAT 1620
Db 1901 ATTCAATTTTATAGCAACCAATTTGTTAAACTCACTGTGATCAATTTTATATCAT 1960
QY 1621 GCAAAATATGTTTAAATTAATGAAATGTAAT 1655
Db 1961 GCAAAATATGTTTAAATTAATGAAATGTAAT 1995

RESULT 14

US-08-999-811-3
; Sequence 3, Application US/08999811
; Patent No. 5932540
; GENERAL INFORMATION:
; APPLICANT: HU JING-SHAN
; APPLICANT: ROSEN, CRAIG A.
; APPLICANT: CAO, LIANG
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
; STREET: 1100 NEW YORK AVENUE
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMEUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patencin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/999,811
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/207,550
; FILING DATE: 8-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/465,968
; FILING DATE: 06-JUN-1995

ATTORNEY/AGENT INFORMATION:
; NAME: MARKOWICZ, KAREN R.
; REGISTRATION NUMBER: 36,351
; REFERENCE/DOCKET NUMBER: 1488.1000004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)371-2600
; TELEFAX: (202)371-2540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1526 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 71..142
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 143..1120
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 71..1120
US-08-999-811-3

Query Match 91.2%; Score 1526; DB 2; Length 1526;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1526; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 149 CGAGGCCACGGCTTATCCAAAGCAAGATCTGGAGGAGCACTTACGGTCTGTGTCAGTGT 208
Db 1 CGAGGCCACGGCTTATCCAAAGCAAGATCTGGAGGAGCACTTACGGTCTGTGTCAGTGT 60
QY 209 AGATGAACCTCATGACTTACTCTACCCAGATAATTTGAAATATGTACAAAGTGTGAGTAA 268
Db 61 AGATGAACCTCATGACTTACTCTACCCAGATAATTTGAAATATGTACAAAGTGTGAGTAA 120
QY 269 GAAAGGAGGCTGGCAACATACAGAGAACAGGCAACCTCAACTCAAGGAGCAAGAGAC 328
Db 121 GAAAGGAGGCTGGCAACATACAGAGAACAGGCAACCTCAACTCAAGGAGCAAGAGAC 180
QY 329 TATAAATTTTCTGCAGCAGCATTTATATACAGAGATCTTGAAAGTATTCATATGAGTG 388
Db 181 TATAAATTTTCTGCAGCAGCATTTATATACAGAGATCTTGAAAGTATTCATATGAGTG 240
QY 389 GAGAAAGACTCAATGCAATGCCACGGAGGCTGTATATAGATGTGGGAAGAGTTTGGAGT 448
Db 241 GAGAAAGACTCAATGCAATGCCACGGAGGCTGTATATAGATGTGGGAAGAGTTTGGAGT 300
QY 449 CGCAGCAAAACACCTTCTTTAAACCTCCATGTGTGCTGCTACAGATGTGGGGTGTGCTG 508
Db 301 CGCAGCAAAACACCTTCTTTAAACCTCCATGTGTGCTGCTACAGATGTGGGGTGTGCTG 360
QY 509 CAATAGTGAGGGGCTGCAGTGCATGAACACAGCAGCTACTCTCAGCAAGACGTTATT 568
Db 361 CAATAGTGAGGGGCTGCAGTGCATGAACACAGCAGCTACTCTCAGCAAGACGTTATT 420
QY 569 TGAATTTACAGTGCCTCTCTCTCAAGGCCCAACCAAGTAACTCAGTTTGGCCAAATCA 628
Db 421 TGAATTTACAGTGCCTCTCTCTCAAGGCCCAACCAAGTAACTCAGTTTGGCCAAATCA 480
QY 629 CACTTCTCTGCGATGCATGCTTAACTGGATGTTTACAGACAGTTTCATTCATTATTAG 688
Db 481 CACTTCTCTGCGATGCATGCTTAACTGGATGTTTACAGACAGTTTCATTCATTATTAG 540
QY 689 ACCTTCTCTGCGAGCAACACTTACCACAGTGTGAGGAGCGAAACAAAGACCTGCCCAACAA 748
Db 541 ACCTTCTCTGCGAGCAACACTTACCACAGTGTGAGGAGCGAAACAAAGACCTGCCCAACAA 600
QY 749 TTACATGTGGAATTAATCAGATCTGAGATGCTGGCTCAGAGATTTTATGTTTCTC 808
Db 601 TTACATGTGGAATTAATCAGATCTGAGATGCTGGCTCAGAGATTTTATGTTTCTC 660

QY 809 GGATGCTGGAGATGATCTCAACAGATGATTCATCATCTGTGGACCAACAAAGAGCT 868
 Db 661 GGATGCTGGAGATGATCTCAACAGATGATTCATCATCTGTGGACCAACAAAGAGCT 720
 QY 869 GGATGAGAGACCTGTGAGTGTCTGAGAGCGGGCTTCGGCTCGCAGCTGTGGACC 928
 Db 721 GGATGAGAGACCTGTGAGTGTCTGAGAGCGGGCTTCGGCTCGCAGCTGTGGACC 780
 QY 929 CCACAAAGATAGACAGAACTCATGCGCATGTGTCTGTAAACAACTCTTCCCGAG 988
 Db 781 CCACAAAGATAGACAGAACTCATGCGCATGTGTCTGTAAACAACTCTTCCCGAG 840
 QY 989 CCAATGTGGGCCCAACCGAGATTTGATGAACACATGCCAGTGTGTATGAAGAAC 1048
 Db 841 CCAATGTGGGCCCAACCGAGATTTGATGAACACATGCCAGTGTGTATGAAGAAC 900
 QY 1049 CTGCCCCAGAAATCAACCCCTAAATCTCTGGAAATGTCCCTGTGAATGTACAGAAAGTCC 1108
 Db 901 CTGCCCCAGAAATCAACCCCTAAATCTCTGGAAATGTCCCTGTGAATGTACAGAAAGTCC 960
 QY 1109 ACAGAAATGCTTGTAAAGGAAGAGTTCACACCAACATGCAGCTGTACAGACG 1168
 Db 961 ACAGAAATGCTTGTAAAGGAAGAGTTCACACCAACATGCAGCTGTACAGACG 1020
 QY 1169 GCCATGTACGAACCGCCAGAGGCTTGTGAGCCAGGATTTTATATAGTGAAGAAGTGTG 1228
 Db 1021 GCCATGTACGAACCGCCAGAGGCTTGTGAGCCAGGATTTTATATAGTGAAGAAGTGTG 1080
 QY 1229 TCGTGTGCTCCCTTCATATGCGAAGACACAAATAGAGTGAATGTACTGTTTCCA 1288
 Db 1081 TCGTGTGCTCCCTTCATATGCGAAGACACAAATAGAGTGAATGTACTGTTTCCA 1140
 QY 1289 GTTCATCGATTTCTTATATGAAATCTGTGTCAGTGAAGTGTCTGTGAAACAGA 1348
 Db 1141 GTTCATCGATTTCTTATATGAAATCTGTGTCAGTGAAGTGTCTGTGAAACAGA 1200
 QY 1349 GAGACCTTGTGGTCCATGCTAAACAAAGACAAAGTCTGTTCTCTGAACCATGTGGA 1408
 Db 1201 GAGACCTTGTGGTCCATGCTAAACAAAGACAAAGTCTGTTCTCTGAACCATGTGGA 1260
 QY 1409 TAATTTACAGAAATGAGTGGCTCATCTGAAAGGCTCTGTAAAGACTGGTTTT 1468
 Db 1261 TAATTTACAGAAATGAGTGGCTCATCTGAAAGGCTCTGTAAAGACTGGTTTT 1320
 QY 1469 CTGCCAATGACCAACAGCCAGATTTTCTCTGTGATTTCTTAAAGAAAGTACTATA 1528
 Db 1321 CTGCCAATGACCAACAGCCAGATTTTCTCTGTGATTTCTTAAAGAAAGTACTATA 1380
 QY 1529 TAATTTATTCACATAAAATATTTGTTCTGCAATCTATTTATAGCAACAAATTTGT 1588
 Db 1381 TAATTTATTCACATAAAATATTTGTTCTGCAATCTATTTATAGCAACAAATTTGT 1440
 QY 1589 AAACTCACTGTGATCAATATTTTATATCATGCAAAATATGTTTAAATATAATGAAA 1648
 Db 1441 AAACTCACTGTGATCAATATTTTATATCATGCAAAATATGTTTAAATATAATGAAA 1500
 QY 1649 TGTATTTATAAAAAA 1674
 Db 1501 TGTATTTATAAAAAA 1526

RESULT 15

US-09-042-105-3

; Sequence 3, Application US/09042105

; Patent No. 6040157

; GENERAL INFORMATION:

; APPLICANT: HU, JING-SHAN

; APPLICANT: ROSEN, CRAIG A.

; APPLICANT: CAO, LIANG

; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2

; NUMBER OF SEQUENCES: 35

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX

STREET: 1100 NEW YORK AVENUE
 CITY: WASHINGTON
 STATE: DC
 COUNTRY: USA
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/042,105
 FILING DATE: HEREWITH
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/207,550
 FILING DATE: 8-MAR-1994
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/465,968
 FILING DATE: 06-JUN-1995
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: TO BE ASSIGNED
 FILING DATE: 24-DEC-1997
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: ERIC K. STEFFE
 REGISTRATION NUMBER: 36,688
 REFERENCE/DOCKET NUMBER: 1488.1000003/EKS
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)371-2600
 TELEFAX: (202)371-2540
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1526 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: CDNA
 FEATURE:
 NAME/KEY: sig_peptide
 LOCATION: 71..142
 FEATURE:
 NAME/KEY: mat_peptide
 LOCATION: 143..1120
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 71..1120
 US-09-042-105-3

Query Match 91.2%; Score 1526; DB 3; Length 1526;

Best Local Similarity 100.0%; Pred. No. 0;

Mismatches 0; Indels 0; Gaps 0;

Matches 1526; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 149 CGAGGCCACGGCTTATGCAAGCAAGATCTGGAGGAGCAGTTACGGTCTGTGTCAGGTG 208

Db 1 CGAGGCCACGGCTTATGCAAGCAAGATCTGGAGGAGCAGTTACGGTCTGTGTCAGGTG 60

QY 209 AGATGAACCTCATGACTGTACTCTACCCAGAAATATTGAAAAATGTACAAAGTGTCAAG 268

Db 61 AGATGAACCTCATGACTGTACTCTACCCAGAAATATTGAAAAATGTACAAAGTGTCAAG 120

QY 269 GAAAGAGGCTGGCAACATACAGAGAACAGCCCACTCACTCAAGGACAGAGAGAC 328

Db 121 GAAAGAGGCTGGCAACATACAGAGAACAGCCCACTCACTCAAGGACAGAGAGAC 180

QY 329 TATAAAATTTGCTGAGCAGACATTATAATACAGAGATCTTGAAGATTTGATAATGAGTG 388

Db 181 TATAAAATTTGCTGAGCAGACATTATAATACAGAGATCTTGAAGATTTGATAATGAGTG 240

QY 389 GAGAAAGACTCAATGCATGCCACCGGAGGTGTATAGATGTGGGAGAGGAGTTGGAGT 448

Db 389 GAGAAAGACTCAATGCATGCCACCGGAGGTGTATAGATGTGGGAGAGGAGTTGGAGT 448

Db 241 CAGAAAGACTCAATGCATGCCACGGAGGTGCTGTATAGATGTGGGAAGGAGTTTGGAGT 300
 Qy 449 CGCAGCAAAACACCTCTTTAAACCTCCATGTGTGCTCGGTCTACAGATGTGGGGTGTGCTG 508
 Db 301 CGCAGCAAAACACCTCTTTAAACCTCCATGTGTGCTCGGTCTACAGATGTGGGGTGTGCTG 360
 Qy 509 CAATAGTGAAGGGGTGAGTGCATGAACACACGACGAGCTTACCTCAGCAAGAGCTTTATT 568
 Db 361 CAATAGTGAAGGGGTGAGTGCATGAACACACGACGAGCTTACCTCAGCAAGAGCTTTATT 420
 Qy 569 TGAATTTACGTGCTCTCTCTCAAGGCCCAACACAGTAACAATCAGTTTTCCTCAATCA 628
 Db 421 TGAATTTACGTGCTCTCTCTCAAGGCCCAACACAGTAACAATCAGTTTTCCTCAATCA 480
 Qy 629 CACTTCCTGCGAGTGTCTTAACCTGAGTGTATACAGACAAGTTTCATTTCCATTTATAG 688
 Db 481 CACTTCCTGCGAGTGTCTTAACCTGAGTGTATACAGACAAGTTTCATTTCCATTTATAG 540
 Qy 689 ACGTTCCCTGCCAGCAACACTTACACAGTGTGAGCGAGCGAACAAGACCTGCCCCACCAA 748
 Db 541 ACGTTCCCTGCCAGCAACACTTACACAGTGTGAGCGAGCGAACAAGACCTGCCCCACCAA 600
 Qy 749 TTACATGTGGAAATTAATCACATCTCCAGATGCTGGCTCAGGAAGATTTTATGTTTTCCTC 808
 Db 601 TTACATGTGGAAATTAATCACATCTCCAGATGCTGGCTCAGGAAGATTTTATGTTTTCCTC 660
 Qy 809 GGATGCTGGAGATGACTCAACAGATGGATTCCATGACATCTGTGGACCAACAAGGAGCT 868
 Db 661 GGATGCTGGAGATGACTCAACAGATGGATTCCATGACATCTGTGGACCAACAAGGAGCT 720
 Qy 869 GGATGAAGAGACCTGTGAGTGTCTGAGAGCGGGCTTCGGCTGCCAGCTGTGGAAC 928
 Db 721 GGATGAAGAGACCTGTGAGTGTCTGAGAGCGGGCTTCGGCTGCCAGCTGTGGAAC 780
 Qy 929 CCACAAAGAACTAGACAGAACTCATGCTGAGTGTCTGTAAACCAAACTCTTCCCCAG 988
 Db 781 CCACAAAGAACTAGACAGAACTCATGCTGAGTGTCTGTAAACCAAACTCTTCCCCAG 840
 Qy 989 CCAATGTGGGCCCAACCGAGAAATTTGATGAACACATGCCAGTGTGTATGTAAAGAAC 1048
 Db 841 CCAATGTGGGCCCAACCGAGAAATTTGATGAACACATGCCAGTGTGTATGTAAAGAAC 900
 Qy 1049 CTGCCCCAGAAATCAACCCCTAAATCTGGAAATGTGCTGTAATGTACAGAAAGTCC 1108
 Db 901 CTGCCCCAGAAATCAACCCCTAAATCTGGAAATGTGCTGTAATGTACAGAAAGTCC 960
 Qy 1109 ACAGAAATGCTGTGTAAAGAAAGAAAGTTCCACCAACCAACATGCGAGCTGTACAGACG 1168
 Db 961 ACAGAAATGCTGTGTAAAGAAAGAAAGTTCCACCAACCAACATGCGAGCTGTACAGACG 1020
 Qy 1169 GCCATGTACGAAACCGCCAGAGGCTTGTGAGCCAGGATTTTCATATAGTGAAGAGTGTG 1228
 Db 1021 GCCATGTACGAAACCGCCAGAGGCTTGTGAGCCAGGATTTTCATATAGTGAAGAGTGTG 1080
 Qy 1229 TCGTTGTGTCCTTTCATATTGGCAAGACCAACATGAGCTTAAGATTGTACTGTGTTTCCA 1288
 Db 1081 TCGTTGTGTCCTTTCATATTGGCAAGACCAACATGAGCTTAAGATTGTACTGTGTTTCCA 1140
 Qy 1289 GTTCATGCAATTTCTATATGAAAAAATGTTGTTGCCACAGTGAAGTGTCTGTGAACAGA 1348
 Db 1141 GTTCATGCAATTTCTATATGAAAAAATGTTGTTGCCACAGTGAAGTGTCTGTGAACAGA 1200
 Qy 1349 GAGACCTTGTGGGTCCATGTAACAGACAAAGTCTGCTTTCCCTGAACCATGTGGA 1408
 Db 1201 GAGACCTTGTGGGTCCATGTAACAGACAAAGTCTGCTTTCCCTGAACCATGTGGA 1260
 Qy 1409 TAACTTTACAGAAATGAGCTGAGCTCATCTGCAAAAAGGCTCTTGTAAAGACTGGTTTT 1468
 Db 1261 TAACTTTACAGAAATGAGCTGAGCTCATCTGCAAAAAGGCTCTTGTAAAGACTGGTTTT 1320
 Qy 1469 CTGCCAATGACCAACAGCCAGATTTTCTCTGTGATTTCTTTAAAGAAATGACTATA 1528
 Db 1321 CTGCCAATGACCAACAGCCAGATTTTCTCTGTGATTTCTTTAAAGAAATGACTATA 1380

Search completed: February 25, 2004, 22:07:19
 Job time : 152 secs

Qy 1529 TAAATTTATTTCCACTAAAAAATATTTCTTCTGCAATTCATTTTATAGCAACCAACAATTCGT 1588
 Db 1381 TAAATTTATTTCCACTAAAAAATATTTCTTCTGCAATTCATTTTATAGCAACCAACAATTCGT 1440
 Qy 1589 AAAAATCACTGTGATCAATATTTTATATCATGCAAAAATATGTTTAAAAATAAAAATGAAAA 1648
 Db 1441 AAAAATCACTGTGATCAATATTTTATATCATGCAAAAATATGTTTAAAAATAAAAATGAAAA 1500
 Qy 1649 TTGTATTTATAAAAAAATAAAAAA 1674
 Db 1501 TTGTATTTATAAAAAAATAAAAAA 1526

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2004, 19:10:43 ; Search time 621 Seconds

(without alignments)
9439-215 Million cell updates/sec

Title: US-09-921-143-1

Perfect score: 1674

Sequence: 1 gctctccaccatgcactcg.....ttataaaaaaaaaaaaaa 1674

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2308684 seqs, 1750822206 residues

Total number of hits satisfying chosen parameters: 4617368

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US05_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1674	100.0	1674	10	US-09-935-726-1 Sequence 1, Appli
2	1674	100.0	1674	11	US-09-921-143-1 Sequence 1, Appli
3	1674	100.0	1674	13	US-10-127-551-1 Sequence 1, Appli
4	1674	100.0	1674	14	US-10-084-488-1 Sequence 1, Appli
5	1674	100.0	1674	14	US-10-120-398-1 Sequence 1, Appli
6	1674	100.0	1674	14	US-10-120-414-1 Sequence 1, Appli
7	1674	100.0	1674	14	US-10-120-377-1 Sequence 1, Appli
8	1664.6	99.4	2031	14	US-10-346-802-1 Sequence 1, Appli
9	1664	99.4	1786	14	US-10-084-817-98 Sequence 98, Appli
10	1663	99.3	2031	14	US-10-346-802-2 Sequence 21, Appli
11	1651.8	98.7	1997	9	US-09-795-006A-21 Sequence 3, Appli
12	1651.8	98.7	1997	10	US-09-375-248-3 Sequence 7, Appli
13	1651.8	98.7	1997	14	US-10-201-386-7 Sequence 23, Appli
14	1651.8	98.7	1997	14	US-10-262-538-23 Sequence 3, Appli
15	1651.8	98.7	2015	14	US-10-081-126-3 Sequence 3, Appli

16	1651.8	98.7	2015	15	US-10-173-718-4	Sequence 4, Appli
17	1526	91.2	1526	11	US-09-921-143-3	Sequence 3, Appli
18	1526	91.2	1526	14	US-10-084-488-3	Sequence 3, Appli
19	1520.2	90.8	1525	10	US-09-935-726-3	Sequence 1, Appli
20	1520.2	90.8	1525	13	US-10-060-523-1	Sequence 3, Appli
21	1520.2	90.8	1525	14	US-10-120-398-3	Sequence 3, Appli
22	1520.2	90.8	1525	14	US-10-120-414-3	Sequence 3, Appli
23	1520.2	90.8	1525	14	US-10-120-377-3	Sequence 3, Appli
24	1266.6	75.7	5283	11	US-09-921-143-36	Sequence 36, Appli
25	1182.2	70.6	1836	14	US-10-201-386-10	Sequence 10, Appli
26	804.2	48.0	1741	14	US-10-201-386-12	Sequence 12, Appli
27	486.4	29.1	59001	15	US-10-173-718-13	Sequence 13, Appli
28	298	17.8	239	14	US-10-346-802-6	Sequence 6, Appli
29	278.2	16.6	337	9	US-09-795-006A-48	Sequence 48, Appli
30	273.4	16.3	409	15	US-10-173-718-12	Sequence 12, Appli
31	231.8	13.8	331	9	US-09-795-006A-156	Sequence 156, App
32	227	13.6	331	9	US-09-795-006A-158	Sequence 158, App
33	212.6	12.7	331	9	US-09-795-006A-154	Sequence 154, App
34	203	12.1	337	9	US-09-795-006A-112	Sequence 112, App
35	195	11.6	337	9	US-09-795-006A-108	Sequence 108, App
36	194	11.6	328	9	US-09-795-006A-76	Sequence 76, Appli
37	188	11.2	332	9	US-09-795-006A-80	Sequence 80, Appli
38	187.4	11.2	379	15	US-10-242-535A-34545	Sequence 34545, A
39	175.8	10.5	328	9	US-09-795-006A-168	Sequence 168, App
40	174.2	10.4	337	9	US-09-795-006A-110	Sequence 110, App
41	173.2	10.3	334	9	US-09-795-006A-78	Sequence 78, Appli
42	169.4	10.1	337	9	US-09-795-006A-104	Sequence 104, App
43	168.4	10.1	334	9	US-09-795-006A-72	Sequence 72, Appli
44	166.2	9.9	337	9	US-09-795-006A-106	Sequence 106, App
45	165.2	9.9	334	9	US-09-795-006A-44	Sequence 44, Appli

ALIGNMENTS

RESULT 1

US-09-935-726-1
; Sequence 1, Application US/09935726
; Publication No. US20030008357A1
; GENERAL INFORMATION:
; APPLICANT: Hu, Jin-Shan
; APPLICANT: Craig, Rosen
; APPLICANT: Cao, Liang
; TITLE OF INVENTION: Vascular Endothelial Growth Factor-2
; FILE REFERENCE: PF112P3D1C1
; CURRENT APPLICATION NUMBER: US/09/935,726
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 09/438,538
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 1
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (12)..(80)
; NAME/KEY: CDS
; LOCATION: (12)..(1268)
; NAME/KEY: mat_peptide
; LOCATION: (81)..()
US-09-935-726-1

Query Match 100.0%; Score 1674; DB 10; Length 1674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GTCTTCCACCATGCATCGCTCGGCTTCTTCTCTGCGGTCTTCTGCTGCGCTG	60
Db	1	GTCTTCCACCATGCATCGCTCGGCTTCTTCTCTGCGGTCTTCTGCTGCGCTG	60
Qy	61	CGCTGTCCCGGCTCTCTCGGAGGCGCGCGCGCGCTTCTGAGTCCGACTCG	120

Db 61 CGTGTCTCCGGGTCCTTCGAGGCGCCCGCGCGCGCTTCGAGTCCGGACTCG 120
Qy 121 ACCTCTCGGACCGGAGCCCGAGCGCGGCGAGCCACGGCTTATGCAAGCAAGATCTGG 180
Db 121 ACCTCTCGGACCGGAGCCCGAGCGCGGCGAGCCACGGCTTATGCAAGCAAGATCTGG 180
Qy 181 AGGAGCAGTTACGGTCTGTCTCCAGTGTAGATGAATCATGACTGTACTCTACCCAGAAT 240
Db 181 AGGAGCAGTTACGGTCTGTCTCCAGTGTAGATGAATCATGACTGTACTCTACCCAGAAT 240
Qy 241 ATTGGAAAAATGTACAAGTGTGAGTAAAGGAGGCTGGCAACATAACAGAGAACAG 300
Db 241 ATTGGAAAAATGTACAAGTGTGAGTAAAGGAGGCTGGCAACATAACAGAGAACAG 300
Qy 301 CCAACCTCACTCAAGAGAGAGAGACTATAAAATTTGCTGCAGACACATTAATATACAG 360
Db 301 CCAACCTCACTCAAGAGAGAGAGACTATAAAATTTGCTGCAGACACATTAATATACAG 360
Qy 361 AGATCTTTGAAAAGTATTGATAATGAGTGGAGAAAGACTCAATGCTGCCACGGGAGTGT 420
Db 361 AGATCTTTGAAAAGTATTGATAATGAGTGGAGAAAGACTCAATGCTGCCACGGGAGTGT 420
Qy 421 GTATAGATGGGGAAGAGGTTGGAGTGCAGCAAAACACCTTTTAAACCTCCATGTG 480
Db 421 GTATAGATGGGGAAGAGGTTGGAGTGCAGCAAAACACCTTTTAAACCTCCATGTG 480
Qy 481 TGTCCGCTACAGATGGGGGTTGCTGCAATAGTGAAGGGCTGCAGTGCATGAACACCA 540
Db 481 TGTCCGCTACAGATGGGGGTTGCTGCAATAGTGAAGGGCTGCAGTGCATGAACACCA 540
Qy 541 GCACGAGTACCTCAGCAAGACGTTATTGAAATTTACAGTGCCTCTCTCAGGCCCCA 600
Db 541 GCACGAGTACCTCAGCAAGACGTTATTGAAATTTACAGTGCCTCTCTCAGGCCCCA 600
Qy 601 ARCCAGTAAACATCAGTTTCCCAATCACACTTCTCCGCGATCATGTCTAACTGGATG 660
Db 601 ARCCAGTAAACATCAGTTTCCCAATCACACTTCTCCGCGATCATGTCTAACTGGATG 660
Qy 661 TTTACAGCAAGTTCATTCCATTTAGACGTTTCCCTGCCAGCAACACTACCACAGTGC 720
Db 661 TTTACAGCAAGTTCATTCCATTTAGACGTTTCCCTGCCAGCAACACTACCACAGTGC 720
Qy 721 AGGAGGGAACAGACCTGCCACCAATTTACATGTGGAAATATCATCTGCGAGATGCC 780
Db 721 AGGAGGGAACAGACCTGCCACCAATTTACATGTGGAAATATCATCTGCGAGATGCC 780
Qy 781 TGGCTCAGGAAGATTTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGGATTC 840
Db 781 TGGCTCAGGAAGATTTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGGATTC 840
Qy 841 ATGACATCTGTGGACCAAAACAAGAGCTGGATGAAGAGACCTGTCTGATGTGCTGAGAG 900
Db 841 ATGACATCTGTGGACCAAAACAAGAGCTGGATGAAGAGACCTGTCTGATGTGCTGAGAG 900
Qy 901 CGGGGCTTCGGCTCGCAGCTGTGGACCCACAAAGAACTAGACAGAACTCATGCCACT 960
Db 901 CGGGGCTTCGGCTCGCAGCTGTGGACCCACAAAGAACTAGACAGAACTCATGCCACT 960
Qy 961 GTGTCTGTAAAAACAAACTCTTCCCGAGCAATGTGGGGCCAAACCGAGAAATTTGATGAAA 1020
Db 961 GTGTCTGTAAAAACAAACTCTTCCCGAGCAATGTGGGGCCAAACCGAGAAATTTGATGAAA 1020
Qy 1021 ACATATGCGAGTGTATGATTAAGAAACCTTCCCGAGAAATCAACCTTAATCTCGGAA 1080
Db 1021 ACATATGCGAGTGTATGATTAAGAAACCTTCCCGAGAAATCAACCTTAATCTCGGAA 1080
Qy 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTGTTTAAAGGAAAGAACTTC 1140
Db 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTGTTTAAAGGAAAGAACTTC 1140
Qy 1141 ACCACCAAAACATGAGCTGTGTAAGAGCGGCATGTACGAAACCGCCAGAGGCTTGTGAGC 1200

Db 1141 ACCACCAAAACATGAGCTGTGTAAGAGCGGCATGTACGAAACCGCCAGAGCTTGTGAGC 1200
Qy 1201 CAGATTTTCATATAGTAGAAGAGTGTGTCGTTGTGTCCTTCATATTTGGCAAGACAC 1260
Db 1201 CAGATTTTCATATAGTAGAAGAGTGTGTCGTTGTGTCCTTCATATTTGGCAAGACAC 1260
Qy 1261 AAATGAGCTAAGATTTGACTGTTTCCAGTTTCATCGATTTTCTATTATGGAAAACTGTGT 1320
Db 1261 AAATGAGCTAAGATTTGACTGTTTCCAGTTTCATCGATTTTCTATTATGGAAAACTGTGT 1320
Qy 1321 TGCCACAGTAGAAGTGTCTGTGAACAGAGAGACCTTTGTGGGTCCATGCTTAACAAAGACA 1380
Db 1321 TGCCACAGTAGAAGTGTCTGTGAACAGAGAGACCTTTGTGGGTCCATGCTTAACAAAGACA 1380
Qy 1381 AAAGTCTGTCTTCTGAAACCATGCGATTAACCTTTCAGAAATGAGCTGGAGCTCATCTG 1440
Db 1381 AAAGTCTGTCTTCTGAAACCATGCGATTAACCTTTCAGAAATGAGCTGGAGCTCATCTG 1440
Qy 1441 CAAAGGCGCTCTGTAAAGACTGGTTTTCTGCCAATGACCAACAGCAAGATTTCTCTC 1500
Db 1441 CAAAGGCGCTCTGTAAAGACTGGTTTTCTGCCAATGACCAACAGCAAGATTTCTCTC 1500
Qy 1501 TTGTGATTTCTTTAAAGAAATGACTATATATTTTATTTTCCACTAAAATATTTGTTCTGC 1560
Db 1501 TTGTGATTTCTTTAAAGAAATGACTATATATTTTATTTTCCACTAAAATATTTGTTCTGC 1560
Qy 1561 ATTCAATTTTATAGCAACAAATTTGTTAAACTCACTGTGATCAATATTTTATATCAT 1620
Db 1561 ATTCAATTTTATAGCAACAAATTTGTTAAACTCACTGTGATCAATATTTTATATCAT 1620
Qy 1621 GCAAAATATGTTTAAATATAAATGAAATTTGATTTTATAAAAAA 1674
Db 1621 GCAAAATATGTTTAAATATAAATGAAATTTGATTTTATAAAAAA 1674

RESULT 2
US-09-921-143-1
; Sequence 1, Application US/09921143
; Publication No. US20030215921A1
; GENERAL INFORMATION:
; APPLICANT: Coleman, Timothy
; TITLE OF INVENTION: Vascular Endothelial Growth Factor-2
; FILE REFERENCE: PFI112P6
; CURRENT APPLICATION NUMBER: US/09/921,143
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/223,276
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1674
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-921-143-1

Query Match 100.0%; Score 1674; DB 11; Length 1674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTCTTCCACCATGCATCGCTGGGCTTCTCTGTGGCGTGTCTCTGTCTGCCCGCTG 60
Db 1 GTCTTCCACCATGCATCGCTGGGCTTCTCTGTGGCGTGTCTCTGTCTGCCCGCTG 60
Qy 61 CGCTGCTCCGGGTCTTCGAGAGGCGCGCGCGCGCGCGCTTCGAGTCCGGACTCG 120
Db 61 CGCTGCTCCGGGTCTTCGAGAGGCGCGCGCGCGCGCGCTTCGAGTCCGGACTCG 120
Qy 121 ACCTCTCGGACCGGAGCCCGAGCGCGGCGAGCCACGGCTTATGCAAGCAAGATCTGG 180
Db 121 ACCTCTCGGACCGGAGCCCGAGCGCGGCGAGCCACGGCTTATGCAAGCAAGATCTGG 180
Qy 181 AGGAGCAGTTACGGTGTGTGTCAGTGTAGATGAATCATGACTGTACTCTACCCAGAAT 240

Db 181 AGGAGCAGTACGGTCTGTGTCAGGTAGATGAACCTCATGACTGTACTCTACCCGAAT 240
Qy 241 ATTGGAATAATGTAAGAAGTGTGCTGAGTGAAGAGAGGCTGCGCAATGAACAGAGAAAGG 300
Db 241 ATTGGAATAATGTAAGAAGTGTGCTGAGTGAAGAGAGGCTGCGCAATGAACAGAGAAAGG 300
Qy 301 CCAACCTCAACTCAAGGACAGAGAGACTATAAATTTGCTGACAGCACTATATATACAG 360
Db 301 CCAACCTCAACTCAAGGACAGAGAGACTATAAATTTGCTGACAGCACTATATATACAG 360
Qy 361 AGATCTTGAAAGTATGTAATGATGAGTGAAGAGAGCTCAATGATGCCAGGAGGTGT 420
Db 361 AGATCTTGAAAGTATGTAATGATGAGTGAAGAGAGCTCAATGATGCCAGGAGGTGT 420
Qy 421 GTATAGATGTGGGAGAGAGTGTGAGTGCAGCAACACACCTCTTTAAACCTCCATGTG 480
Db 421 GTATAGATGTGGGAGAGAGTGTGAGTGCAGCAACACACCTCTTTAAACCTCCATGTG 480
Qy 481 TGTCGGTCTACAGATGTGGGGTGTCTGCAATAGTGAAGGGCTGACGTGATGAACACCA 540
Db 481 TGTCGGTCTACAGATGTGGGGTGTCTGCAATAGTGAAGGGCTGACGTGATGAACACCA 540
Qy 541 GCACGAGTACTGACGAGAGGTATTTGAATTTACAGTGCCTCTCTCAAGGCCCA 600
Db 541 GCACGAGTACTGACGAGAGGTATTTGAATTTACAGTGCCTCTCTCAAGGCCCA 600
Qy 601 AACCACTAACATCAGTTTGGCAATCACACTCTGCGGATGATGCTAAACTGGATG 660
Db 601 AACCACTAACATCAGTTTGGCAATCACACTCTGCGGATGATGCTAAACTGGATG 660
Qy 661 TTTACAGACAAGTTCATTCATTTAGAGCTTCCCTGCGAGAACCACTACCAAGTGC 720
Db 661 TTTACAGACAAGTTCATTCATTTAGAGCTTCCCTGCGAGAACCACTACCAAGTGC 720
Qy 721 AGGAGCGAACAAGACCTGCCCACTTACATGTGGAATTAACATCTGCAGATGCC 780
Db 721 AGGAGCGAACAAGACCTGCCCACTTACATGTGGAATTAACATCTGCAGATGCC 780
Qy 781 TGGCTCAGGAAGATTTATGTTTTCTCGGATGCTGGAGATGACTCAACAGATGGAATCC 840
Db 781 TGGCTCAGGAAGATTTATGTTTTCTCGGATGCTGGAGATGACTCAACAGATGGAATCC 840
Qy 841 ATGACATCTGGACCAACAAGAGCTGGATGAAGAGACCTGTCAGTGTCTGACAG 900
Db 841 ATGACATCTGGACCAACAAGAGCTGGATGAAGAGACCTGTCAGTGTCTGACAG 900
Qy 901 CGGGCTTCGGCTGCCAGCTGTGACCCCAAGAACTAGACAGAACTCATGCCAGT 960
Db 901 CGGGCTTCGGCTGCCAGCTGTGACCCCAAGAACTAGACAGAACTCATGCCAGT 960
Qy 961 GTGCTGTAAAAACAACACTTCCCGAGCCCAATGTGGGGCAACCGAGAAATTTGATGAAA 1020
Db 961 GTGCTGTAAAAACAACACTTCCCGAGCCCAATGTGGGGCAACCGAGAAATTTGATGAAA 1020
Qy 1021 ACACATGCGAGTGTATGTAAGAACTGCGCCCAAGAACTCAACCCCTAAATCCCTGGAA 1080
Db 1021 ACACATGCGAGTGTATGTAAGAACTGCGCCCAAGAACTCAACCCCTAAATCCCTGGAA 1080
Qy 1081 AATGTCCTGCAATGTACAGAAAGTCCACAGAAATGCTTTGTAAGAGAAAGAGTTCC 1140
Db 1081 AATGTCCTGCAATGTACAGAAAGTCCACAGAAATGCTTTGTAAGAGAAAGAGTTCC 1140
Qy 1141 ACCACCAACATGAGCTGTACAGAGCGGCATGTACAGACCGCAGAAAGCTGTGAGC 1200
Db 1141 ACCACCAACATGAGCTGTACAGAGCGGCATGTACAGACCGCAGAAAGCTGTGAGC 1200
Qy 1201 CAGGATTTTATATAGTGAAGAGTGTGCTGTTGTCCTTTCATATTTGGCAAGACCA 1260
Db 1201 CAGGATTTTATATAGTGAAGAGTGTGCTGTTGTCCTTTCATATTTGGCAAGACCA 1260
Qy 1261 AAATGAGCTAAGATGTACTGTTTCCAGTTTCAGATTTTCTATTTGAAAACTGTGT 1320
Db 1261 AAATGAGCTAAGATGTACTGTTTCCAGTTTCAGATTTTCTATTTGAAAACTGTGT 1320

RESULT 3

US-10-127-551-1
; Sequence 1, Application US/10127551
; Publication No. US20020120123A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Vascular Endothelial Growth Factor 2
; FILE REFERENCE: PPI12P1
; CURRENT APPLICATION NUMBER: US/10/127,551
; PRIOR FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/465,968
; PRIOR FILING DATE: EARLIER FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/207,550
; PRIOR FILING DATE: EARLIER FILING DATE: 1994-03-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1674
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (12)..(1268)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (12)..(149)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (150)..(1268)
US-10-127-551-1

Query Match 100.0%; Score 1674; DB 13; Length 1674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTCCCTCCACCACTGACCTGCTGGGCTTCTCTCTGCGCGTCTCTCTGCGCGGCTG 60
Db 1 GTCCCTCCACCACTGACCTGCTGGGCTTCTCTCTGCGCGTCTCTCTGCGCGGCTG 60
Qy 61 CGCTGCTCCCGGGTCTCTGCGAGGCGCGCGCGCGCGCGCGCTTCGAGTCCGAGCTCG 120
Db 61 CGCTGCTCCCGGGTCTCTGCGAGGCGCGCGCGCGCGCGCTTCGAGTCCGAGCTCG 120
Qy 121 ACCTCTCGGACCGGAGCCCGAGCGGGGAGGCCACCGGCTTATGCAAGCAAGATCTGG 180
Db 121 ACCTCTCGGACCGGAGCCCGAGCGGGGAGGCCACCGGCTTATGCAAGCAAGATCTGG 180

QY 181 AGGAGCAGTTACGGTCTGTGTCAGTGTAGATGAACCTCATGCTACTCTACCCAGAAT 240
Db 181 AGGAGCAGTTACGGTCTGTGTCAGTGTAGATGAACCTCATGCTACTCTACCCAGAAT 240
QY 241 ATTGGAATAATGTAAGTGTACAGTGTACAGTGTAGAGGAAGGGCTGGCAACATAACAGAGAACAGG 300
Db 241 ATTGGAATAATGTAAGTGTACAGTGTAGAGGAAGGGCTGGCAACATAACAGAGAACAGG 300
QY 301 CCAACCTCAACTCAAGGACGAAGAGACTATAAAATTTGTCGAGCACATTAATAATACAG 360
Db 301 CCAACCTCAACTCAAGGACGAAGAGACTATAAAATTTGTCGAGCACATTAATAATACAG 360
QY 361 AGATCTTGAAAGTATTGATAATCACTGGAGAAAGACTCAATGATGCCACGGAGGTGT 420
Db 361 AGATCTTGAAAGTATTGATAATCACTGGAGAAAGACTCAATGATGCCACGGAGGTGT 420
QY 421 GTATAGATGTGGGGAAGAGTTTGGAGTCGCGAACAACACCTCTTTAAACCTCCATGTG 480
Db 421 GTATAGATGTGGGGAAGAGTTTGGAGTCGCGAACAACACCTCTTTAAACCTCCATGTG 480
QY 481 TGTCGGTCTACAGATGTGGGGTTGCTGCAATAGTGAGGGGCTGCAGTGATGAACACCA 540
Db 481 TGTCGGTCTACAGATGTGGGGTTGCTGCAATAGTGAGGGGCTGCAGTGATGAACACCA 540
QY 541 GCACGAGTACTCTCAGCAAGACGTTATTGAAATTACAGTGCCTCTCTCAAGGGCCCCA 600
Db 541 GCACGAGTACTCTCAGCAAGACGTTATTGAAATTACAGTGCCTCTCTCAAGGGCCCCA 600
QY 601 AACCAAGTAACATCAGTTTGGCCAAATCACACTTCTGCGGATGCTGTCTAAACTGGATG 660
Db 601 AACCAAGTAACATCAGTTTGGCCAAATCACACTTCTGCGGATGCTGTCTAAACTGGATG 660
QY 661 TTTACAGACAGTTTCATTCCATTATTAGAGTTCCTGCGCAGCAACACTACACAGTGTC 720
Db 661 TTTACAGACAGTTTCATTCCATTATTAGAGTTCCTGCGCAGCAACACTACACAGTGTC 720
QY 721 AGGCAGCAACAGACCTGCCCAACCAATTAATGAGTGAATAATCAATCTGCAGATGCC 780
Db 721 AGGCAGCAACAGACCTGCCCAACCAATTAATGAGTGAATAATCAATCTGCAGATGCC 780
QY 781 TGGCTCAGGAGATTTTATGTTTCTCGATGCTGAGATGACTCAACAGATGGATTC 840
Db 781 TGGCTCAGGAGATTTTATGTTTCTCGATGCTGAGATGACTCAACAGATGGATTC 840
QY 841 ATGACATCTGTGGACCAACAAAGAGCTGGATGAAGAGACTGTGAGTGTCTGCAGAG 900
Db 841 ATGACATCTGTGGACCAACAAAGAGCTGGATGAAGAGACTGTGAGTGTCTGCAGAG 900
QY 901 CGGGCTTCGGCTTCGAGCTGTGGACCCCAACAAAGAACTAGACAGAACTCATGCCAGT 960
Db 901 CGGGCTTCGGCTTCGAGCTGTGGACCCCAACAAAGAACTAGACAGAACTCATGCCAGT 960
QY 961 GTGCTGTAAAAACAACACTCTCCCGAGCAATGTGGGGCAACCGAGAAATTTGATGAAA 1020
Db 961 GTGCTGTAAAAACAACACTCTCCCGAGCAATGTGGGGCAACCGAGAAATTTGATGAAA 1020
QY 1021 ACATGCGCAGTGTGTATGTAAGAAACCTGCGCCAGAAATCAACCCCTAAATCTCGAA 1080
Db 1021 ACATGCGCAGTGTGTATGTAAGAAACCTGCGCCAGAAATCAACCCCTAAATCTCGAA 1080
QY 1081 AATGCTGCTGGAATGTACAGAAAGTCCACAGAAATGCTTTGTAAGGAAGAAAGTTCC 1140
Db 1081 AATGCTGCTGGAATGTACAGAAAGTCCACAGAAATGCTTTGTAAGGAAGAAAGTTCC 1140
QY 1141 ACCACCAACATGAGCTGTGTACAGACGGCCATGTACGAAACCGCCAGAGGCTTGTGAGC 1200
Db 1141 ACCACCAACATGAGCTGTGTACAGACGGCCATGTACGAAACCGCCAGAGGCTTGTGAGC 1200
QY 1201 CAGATTTTCATATAGTGAAGAGTGTGCTGCTGTCCTTCATATTGTCGAAGACCCAC 1260
Db 1201 CAGATTTTCATATAGTGAAGAGTGTGCTGCTGTCCTTCATATTGTCGAAGACCCAC 1260
QY 1261 AAATGAGCTAGATTTGACTGTTTCCAGTTTCATCGATTTTCTATTATGGAAGAACTGTGT 1320

RESULT 4

US-10-084-488-1
; Sequence 1, Application US/10084488
; Publication No. US20030028007A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/084,488
; FILING DATE: 28-Feb-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/623,725
; FILING DATE: 07-Sep-2000
; APPLICATION NUMBER: US/09/042,105
; FILING DATE: 13-MAR-1998
; APPLICATION NUMBER: US/09/107,997
; FILING DATE: 30-JUN-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: MICHELE M. WALES
; REGISTRATION NUMBER: 43,975
; REFERENCE/DOCKET NUMBER: PF112PCT3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301)309-8504
; TELEFAX: (301)309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1674 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA

FEATURE: NAME/KEY: sig_peptide
LOCATION: 12...80
FEATURE: NAME/KEY: mat_peptide
LOCATION: 81...1268
FEATURE: NAME/KEY: CDS
LOCATION: 12...1268
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-084-488-1

Query Match 100.0%; Score 1674; DB 14; Length 1674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCTTCCACCATGCACTCGCTGGGCTTCTTCTGTGGCGTGTCTCTGTCTCGCGGTG 60
DB 1 GTCTTCCACCATGCACTCGCTGGGCTTCTTCTGTGGCGTGTCTCTGTCTCGCGGTG 60

QY 61 CGTGTCTCCGGTCTCTCGAGGGCCCGCGCGCGCGCGCGCGCTTCGAGTCCGACTCG 120
DB 61 CGTGTCTCCGGTCTCTCGAGGGCCCGCGCGCGCGCGCGCGCTTCGAGTCCGACTCG 120

QY 121 ACCTCTCGGACGCGGAGCCCGCGCGGCGAGGCCACGCGCTTATGCAAGCAAGATCTGG 180
DB 121 ACCTCTCGGACGCGGAGCCCGCGCGGCGAGGCCACGCGCTTATGCAAGCAAGATCTGG 180

QY 181 AGGAGCAGTACGCTGTGTCTGCTAGTGTAGTGAATCTATGACTGTCTTACCCAGAT 240
DB 181 AGGAGCAGTACGCTGTGTCTGCTAGTGTAGTGAATCTATGACTGTCTTACCCAGAT 240

QY 241 ATTGGAATGTACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB 241 ATTGGAATGTACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300

QY 301 CCACTCTCAACTCAAGGACAGAGACTATATAATTTGCTCGACGACATTAATATACAG 360
DB 301 CCACTCTCAACTCAAGGACAGAGACTATATAATTTGCTCGACGACATTAATATACAG 360

QY 361 AGATCTTGAAGTATTGATATAGTGGAGAGAGACTCAATGCGATGCCGAGGAGT 420
DB 361 AGATCTTGAAGTATTGATATAGTGGAGAGAGACTCAATGCGATGCCGAGGAGT 420

QY 421 GTATAGATGTGGGAGAGGTTTGGAGTCCGACAAACACCTTCTTTAAACCTCCATGTG 480
DB 421 GTATAGATGTGGGAGAGGTTTGGAGTCCGACAAACACCTTCTTTAAACCTCCATGTG 480

QY 481 TGTCCGTCTACAGATGTGGGTTGCTGCAATAGTGGAGGCTGAGTGCATGAACACCA 540
DB 481 TGTCCGTCTACAGATGTGGGTTGCTGCAATAGTGGAGGCTGAGTGCATGAACACCA 540

QY 541 GCACGAGTACCTCAGCAAGAGCTTATTTGAAATTACAGTGTCTCTCAAGGCCCA 600
DB 541 GCACGAGTACCTCAGCAAGAGCTTATTTGAAATTACAGTGTCTCTCAAGGCCCA 600

QY 601 AACCACTAAATCAGTTTGGCAATCACTTCTGCGGATGCAATGCTAACTGATG 660
DB 601 AACCACTAAATCAGTTTGGCAATCACTTCTGCGGATGCAATGCTAACTGATG 660

QY 661 TTTACAGACAGTTCATTCATATTAGAGTTCCTGCGGACCACTACACAGTGTG 720
DB 661 TTTACAGACAGTTCATTCATATTAGAGTTCCTGCGGACCACTACACAGTGTG 720

QY 721 AGGACGCAACAGACTGCCCCCAATTAATGATGGAATATCACTCTCAGATGCC 780
DB 721 AGGACGCAACAGACTGCCCCCAATTAATGATGGAATATCACTCTCAGATGCC 780

QY 781 TGCTCAGGAGATTTATGTTTTCTCGAGTGTGAGTCACTCAACAGATGATCC 840
DB 781 TGCTCAGGAGATTTATGTTTTCTCGAGTGTGAGTCACTCAACAGATGATCC 840

QY 841 ATGACATCTGTGACCAACAGAGCTGGATGAAGAGACCTGTCTAGTGTCTGAGAG 900

DB 841 ATGACATCTGTGACCAACAGAGAGCTGGATGAAGAGACCTGTCTAGTGTCTGAGAG 900
QY 901 CGGGGCTTCGGCTGCGAGCTGTGGAGCCCAACAGAACTAGACAGAAATCTATGCCAGT 960
DB 901 CGGGGCTTCGGCTGCGAGCTGTGGAGCCCAACAGAACTAGACAGAAATCTATGCCAGT 960

QY 961 GTGTCTGTAATAACAAACTCTTCCAGCCCAATGTGGGCGCAACCGAGATTTGATGAAA 1020
DB 961 GTGTCTGTAATAACAAACTCTTCCAGCCCAATGTGGGCGCAACCGAGATTTGATGAAA 1020

QY 1021 ACACATGCCAGTGTGTATGTAAAGAACCTGCCCCAGAAATCAACCCCTAAATCTGGAA 1080
DB 1021 ACACATGCCAGTGTGTATGTAAAGAACCTGCCCCAGAAATCAACCCCTAAATCTGGAA 1080

QY 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTTAAAGGAAAGAGTTCC 1140
DB 1081 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTTAAAGGAAAGAGTTCC 1140

QY 1141 ACCACCAACATGCGAGCTGTTACAGAGCGGCATGTAGAACCCGCAAGAGCTTGTGAGC 1200
DB 1141 ACCACCAACATGCGAGCTGTTACAGAGCGGCATGTAGAACCCGCAAGAGCTTGTGAGC 1200

QY 1201 CAGGATTTTCAATAGTGAAGAGTGTGCTGTGCTTCCAGTTTCTATATGSCAAAGACCAC 1260
DB 1201 CAGGATTTTCAATAGTGAAGAGTGTGCTGTGCTTCCAGTTTCTATATGSCAAAGACCAC 1260

QY 1261 AAATGAGCTAAGATGCTGCTGCTTCCAGTTTCTGAGTTTCTATATGSCAAAGTTGT 1320
DB 1261 AAATGAGCTAAGATGCTGCTGCTTCCAGTTTCTGAGTTTCTATATGSCAAAGTTGT 1320

QY 1321 TGCCACAGTAGAAGTGTGCTGGAACAGAGAGACCTTGTGGTCCATGTCTAACAAAGACA 1380
DB 1321 TGCCACAGTAGAAGTGTGCTGGAACAGAGAGACCTTGTGGTCCATGTCTAACAAAGACA 1380

QY 1381 AAAGTCTGCTTCTTCCAGTGTGCTGGAACAGTGTGGAACAGTGTGCTGCTGCTGCTG 1440
DB 1381 AAAGTCTGCTTCTTCCAGTGTGCTGGAACAGTGTGGAACAGTGTGCTGCTGCTGCTG 1440

QY 1441 CAAAGGCTCTTGTAAAGACTGCTTCTGCGCAATGACCAACAGCAAGATTTTCTC 1500
DB 1441 CAAAGGCTCTTGTAAAGACTGCTTCTGCGCAATGACCAACAGCAAGATTTTCTC 1500

QY 1501 TTGTGATTTCTTTAAAGAAATGACTATATAATTTTCCACTTAAATAATTTGTTTCTG 1560
DB 1501 TTGTGATTTCTTTAAAGAAATGACTATATAATTTTCCACTTAAATAATTTGTTTCTG 1560

QY 1561 ATTCAATTTTATAGCAACAAATTTGTAATACTCTGATCAATATTTTATATCAT 1620
DB 1561 ATTCAATTTTATAGCAACAAATTTGTAATACTCTGATCAATATTTTATATCAT 1620

QY 1621 GCAAAATATGTTTAAATAAAATGAAATTTGATTTTATAAAAAA 1674
DB 1621 GCAAAATATGTTTAAATAAAATGAAATTTGATTTTATAAAAAA 1674

RESULT 5
US-10-120-398-1
; Sequence 1, Application US/10120398
; Publication No. US20030170786A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig, et al.
; TITLE OF INVENTION: Vascular Endothelial Growth Factor-2
; FILE REFERENCE: Pfl12P7
; CURRENT APPLICATION NUMBER: US/10/120,398
; CURRENT FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/283,408
; PRIOR FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1674
; TYPE: DNA

ORGANISM: Homo sapiens	
FEATURE:	
NAME/KEY: sig_peptide	
LOCATION: (12)..(80)	
OTHER INFORMATION:	
FEATURE:	
NAME/KEY: CDS	
LOCATION: (12)..(1268)	
OTHER INFORMATION:	
FEATURE:	
NAME/KEY: mat_peptide	
LOCATION: (81)..()	
OTHER INFORMATION:	
US-10-120-398-1	
Query Match 100.0%; Score 1674; DB 14; Length 1674;	
Best Local Similarity 100.0%; Pred. No. 0;	
Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 GTGCTTCACATGCACTCGCTGGGCTTCTTCTGTGGGGTCTTCTGTGCGGCTTCTGTGCTGCGCGCTG 60
Db	1 GTGCTTCACATGCACTCGCTGGGCTTCTTCTGTGGGGTCTTCTGTGCGGCTTCTGTGCTGCGCGCTG 60
Qy	61 CGTGCTCCCGGGTCTCGGAGCG 120
Db	61 CGTGCTCCCGGGTCTCGGAGCG 120
Qy	121 ACCTCTCGAGCGGAGCG 180
Db	121 ACCTCTCGAGCGGAGCG 180
Qy	181 AGGAGCAGTTACGGTCTGTGTCAGTGTAGTGAATCTCATGACTGTACTTACCCAGAAAT 240
Db	181 AGGAGCAGTTACGGTCTGTGTCAGTGTAGTGAATCTCATGACTGTACTTACCCAGAAAT 240
Qy	241 ATTGGAAATGTACAACTGTGCTAAGGAAAGAGGCTGGCAACATAACAGAGAACAGG 300
Db	241 ATTGGAAATGTACAACTGTGCTAAGGAAAGAGGCTGGCAACATAACAGAGAACAGG 300
Qy	301 CCAACCTCACTCAAGCAGAGAGAGACTATAAATTTGCTGCGAGCACAATTATAATACAG 360
Db	301 CCAACCTCACTCAAGCAGAGAGAGACTATAAATTTGCTGCGAGCACAATTATAATACAG 360
Qy	361 AGATCTTTGAAAGTATTGATAATGAGTGGAGAAAGACTCAATGATGCGCAGGGAGGTGT 420
Db	361 AGATCTTTGAAAGTATTGATAATGAGTGGAGAAAGACTCAATGATGCGCAGGGAGGTGT 420
Qy	421 GTATAGATGGGGAAGAGTTGGAGTCCGACAAACACCTTCTTTAACTCCATGTG 480
Db	421 GTATAGATGGGGAAGAGTTGGAGTCCGACAAACACCTTCTTTAACTCCATGTG 480
Qy	481 TGTCCGTCTACAGATGTGGGGTGTCTGCAATAGTGGGGCTGCAGTGCATGAACACCA 540
Db	481 TGTCCGTCTACAGATGTGGGGTGTCTGCAATAGTGGGGCTGCAGTGCATGAACACCA 540
Qy	541 GCACGAGTACCTCAGCAAGAGTTTGAATTTAGAGTGCCTCTCTCTCAAGGGCCCCA 600
Db	541 GCACGAGTACCTCAGCAAGAGTTTGAATTTAGAGTGCCTCTCTCTCTCAAGGGCCCCA 600
Qy	601 AACCAAGTAAACATCAGTTTTCGCAATCACAATCTCCCGATGCTGTCTTAACTGGATG 660
Db	601 AACCAAGTAAACATCAGTTTTCGCAATCACAATCTCCCGATGCTGTCTTAACTGGATG 660
Qy	661 TTTACAGCAAGTTTCATTCATTTATAGAGTTCCTCCCGCAGCAACACTACCAAGTGTG 720
Db	661 TTTACAGCAAGTTTCATTCATTTATAGAGTTCCTCCCGCAGCAACACTACCAAGTGTG 720
Qy	721 AGGAGCGGAACAGACCTGCCCCACCAATACATGTGGAATATCATCTGTCAGATGCC 780
Db	721 AGGAGCGGAACAGACCTGCCCCACCAATATCATGTGGAATATCATCTGTCAGATGCC 780
Qy	781 TGGCTCAGGAAGATTTTATGTTTCTCGGATGTGGAGATGACTCAACAGATGGATTCC 840
Db	781 TGGCTCAGGAAGATTTTATGTTTCTCGGATGTGGAGATGACTCAACAGATGGATTCC 840

RESULT 6
US-10-120-414-1
; Sequence 1, Application US/10120414
; Publication No. US20030175274A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig, et al.
; TITLE OF INVENTION: Vascular Endothelial Growth Factor-2
; FILE REFERENCE: PFI12P9
; CURRENT APPLICATION NUMBER: US/10/120,414
; CURRENT FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/283,385
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/350,366
; PRIOR FILING DATE: 2002-01-24

```

; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1674
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (12)..(80)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (12)..(1268)
; OTHER INFORMATION:
; NAME/KEY: mat_peptide
; LOCATION: (81)..(1)
; OTHER INFORMATION:
US-10-120-414-1

Query Match      100.0%; Score 1674; DB 14; Length 1674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  GTCTTCCACCATGCTCGTGGCTTCTCTCTGTGGCGTGTCTCTGTCGCCGCTG 60
DB      1  GTCTTCCACCATGCTCGTGGCTTCTCTCTGTGGCGTGTCTCTGTCGCCGCTG 60

QY      61  CGCTGCTCCCGGTCCTCGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCG 120
DB      61  CGCTGCTCCCGGTCCTCGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCG 120

QY      121  ACCTCTCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCTGG 180
DB      121  ACCTCTCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCTGG 180

QY      181  AGGAGCAGTACGGTCTGTGTCAGTGTAGATGAATCATGACTGTACTTACCCAGAAT 240
DB      181  AGGAGCAGTACGGTCTGTGTCAGTGTAGATGAATCATGACTGTACTTACCCAGAAT 240

QY      241  ATTGAAATATGACAAATGTAGCTAAGGAAAGAGGCTGGCAACATACAGAAACAGG 300
DB      241  ATTGAAATATGACAAATGTAGCTAAGGAAAGAGGCTGGCAACATACAGAAACAGG 300

QY      301  CCAACTCACTCAAGGACAGAGAGACTATAAAATTTGCTGCAGCACTATATATACAG 360
DB      301  CCAACTCACTCAAGGACAGAGAGACTATAAAATTTGCTGCAGCACTATATATACAG 360

QY      361  AGATCTTGAAATATGATATAGTGGAGAAAGACTCAATGCTATGCCAGGAGGTGT 420
DB      361  AGATCTTGAAATATGATATAGTGGAGAAAGACTCAATGCTATGCCAGGAGGTGT 420

QY      421  GTATAGATGTGGGAGAGAGTTTGGAGTGGGAGCAACACCTCTTTAAACCTCCATGTG 480
DB      421  GTATAGATGTGGGAGAGAGTTTGGAGTGGGAGCAACACCTCTTTAAACCTCCATGTG 480

QY      481  TGTCCGTCTACAGATGTGGGGGTGTGCTGCAATAGTGGGGGTGCGAGTGCATGAACCA 540
DB      481  TGTCCGTCTACAGATGTGGGGGTGTGCTGCAATAGTGGGGGTGCGAGTGCATGAACCA 540

QY      541  GCACGAGCTACCTCAGCAGAGAGTATTGAAATTTACAGTGCCTCTCTCAAGGCCCA 600
DB      541  GCACGAGCTACCTCAGCAGAGAGTATTGAAATTTACAGTGCCTCTCTCAAGGCCCA 600

QY      601  ACCAGTACATCAGTTTGGCAATCACACTTCTGCGGATGCTGCTAACTGGATG 660
DB      601  ACCAGTACATCAGTTTGGCAATCACACTTCTGCGGATGCTGCTAACTGGATG 660

QY      661  TTTACAGCAAGTTTCACTATTATAGAGTTTCCCTGCCAGCAACACTACACAGTGTG 720
DB      661  TTTACAGCAAGTTTCACTATTATAGAGTTTCCCTGCCAGCAACACTACACAGTGTG 720

QY      721  AGGCAGGACACAGACTGCCCCACCAATTAATGATGGGATATACATCTCAGATGCC 780
```

```

DB      721  AGGCAGGACACAGACTGCCCCACCAATTAATGATGGGATATACATCTCAGATGCC 780
QY      781  TGGCTCAGGAAGATTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGATTC 840
DB      781  TGGCTCAGGAAGATTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGATTC 840
QY      841  ATGACATCTGTGGACCAACAGAGGCTGGATGAGAGACCTGTGAGTGTCTCCAGAG 900
DB      841  ATGACATCTGTGGACCAACAGAGGCTGGATGAGAGACCTGTGAGTGTCTCCAGAG 900
QY      901  CGGGGCTTCGGCTCGCAGCTGTGGACCCCAAGAGAACTAGACAGAACTCATGCCAGT 960
DB      901  CGGGGCTTCGGCTCGCAGCTGTGGACCCCAAGAGAACTAGACAGAACTCATGCCAGT 960
QY      961  GTGTCTGTAAAAAACAACCTTTTCCAGCAATGTGGGGCCAAACCGAGAAATTTGATGAA 1020
DB      961  GTGTCTGTAAAAAACAACCTTTTCCAGCAATGTGGGGCCAAACCGAGAAATTTGATGAA 1020
QY      1021  ACACATGCCAGTGTGTATGTAAGAACTCTGCCAGAAATCAACCCCTAAATCTGGAA 1080
DB      1021  ACACATGCCAGTGTGTATGTAAGAACTCTGCCAGAAATCAACCCCTAAATCTGGAA 1080
QY      1081  AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGAAAGAGTTCC 1140
DB      1081  AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTAAAGAAAGAGTTCC 1140
QY      1141  ACCACAAACATGTCAGCTGTGTACAGAGCGCCATGTACGAAACCGCAGAGGCTGTGAGC 1200
DB      1141  ACCACAAACATGTCAGCTGTGTACAGAGCGCCATGTACGAAACCGCAGAGGCTGTGAGC 1200
QY      1201  CAGGATTTTCATATAGTGAAGAGTGTGTCGTGTGTGTCCTTTCATATTTGGCAAGCCAC 1260
DB      1201  CAGGATTTTCATATAGTGAAGAGTGTGTCGTGTGTGTCCTTTCATATTTGGCAAGCCAC 1260
QY      1261  AATGAGCTAAGATTTGTTTCCAGTTCATCGATTTCTATTTATGGAAGAACTGTGT 1320
DB      1261  AATGAGCTAAGATTTGTTTCCAGTTCATCGATTTCTATTTATGGAAGAACTGTGT 1320
QY      1321  TGGCAGCTAGAACTCTGTGTGAACAGAGAGACCTTTGGGGTCCATGTACAAAGACA 1380
DB      1321  TGGCAGCTAGAACTCTGTGTGAACAGAGAGACCTTTGGGGTCCATGTACAAAGACA 1380
QY      1381  AAAGTCTGTCTTTCCTGAAACCATGTGATAAATTTACAGAAATGGAGCTCATCTG 1440
DB      1381  AAAGTCTGTCTTTCCTGAAACCATGTGATAAATTTACAGAAATGGAGCTCATCTG 1440
QY      1441  CAAAGGCTCTGTGTAAAGACTGTTTCTGCCAATGACCAACAGCCCAAGATTTTCCTC 1500
DB      1441  CAAAGGCTCTGTGTAAAGACTGTTTCTGCCAATGACCAACAGCCCAAGATTTTCCTC 1500
QY      1501  TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTTAAATATTTTCTGC 1560
DB      1501  TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTTAAATATTTTCTGC 1560
QY      1561  ATTCAATTTTATAGCAACAAATTTGGTAAACTCTGATGATCATATTTTATATCAT 1620
DB      1561  ATTCAATTTTATAGCAACAAATTTGGTAAACTCTGATGATCATATTTTATATCAT 1620
QY      1621  GCAAAATATGTTTAAATAAATGAAATTTGATTTTATAAAAAAATTTTAAAAA 1674
DB      1621  GCAAAATATGTTTAAATAAATGAAATTTGATTTTATAAAAAAATTTTAAAAA 1674
```

RESULT 7
US-10-120-377-1
; Sequence 1, Application US/10120377
; Publication No. US20030176674A1
; GENERAL INFORMATION:
; APPLICANT: Roben, Craig, et al.
; TITLE OF INVENTION: Vascular Endothelial Growth Factor-2
; FILE REFERENCE: PF112P8
; CURRENT APPLICATION NUMBER: US/10/120,377

;; CURRENT FILING DATE: 2002-04-12
;; PRIOR APPLICATION NUMBER: 60/283,391
;; PRIOR FILING DATE: 2001-04-13
;; PRIOR APPLICATION NUMBER: 60/317,600
;; PRIOR FILING DATE: 2001-09-07
;; NUMBER OF SEQ ID NOS: 79
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 1
;; LENGTH: 1674
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: sig_peptide
;; LOCATION: (12)..(80)
;; OTHER INFORMATION:
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (12)..(1268)
;; OTHER INFORMATION:
;; NAME/KEY: mat_peptide
;; LOCATION: (81)..()
;; OTHER INFORMATION:
;; US-10-120-377-1

Query Match 100.0%; Score 1674; DB 14; Length 1674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGCTTCCACATGCACTGCTGGGTTCTCTCTGTGGGTTCTCTGCTGCGCGTG 60
Db 1 GTGCTTCCACATGCACTGCTGGGTTCTCTCTGTGGGTTCTCTGCTGCGCGTG 60
Qy 61 CGTGCTCCCGGTCCTCGAGGCGCCCGCGCGCGCGCTTCGAGTCGGACTCG 120
Db 61 CGTGCTCCCGGTCCTCGAGGCGCCCGCGCGCGCGCTTCGAGTCGGACTCG 120
Qy 121 ACCTCTCGGACGCGGACCGGCGCGGCGGAGCCACGGCTTATGCAAGCAAGATCTGG 180
Db 121 ACCTCTCGGACGCGGACCGGCGCGGCGGAGCCACGGCTTATGCAAGCAAGATCTGG 180
Qy 181 AGGAGCAGTTACGGTCTGTCTCCAGTGTAGATGAACTCATGACTGTACTCTACCCAGAT 240
Db 181 AGGAGCAGTTACGGTCTGTCTCCAGTGTAGATGAACTCATGACTGTACTCTACCCAGAT 240
Qy 241 ATTGGAATAATGTACAAGTGTCAAGTAAAGGAGGCTGGCAACATAACAGAGAACAGG 300
Db 241 ATTGGAATAATGTACAAGTGTCAAGTAAAGGAGGCTGGCAACATAACAGAGAACAGG 300
Qy 301 CCAACCTCAACTAAGGACAGAGAGACTATAAAATTTGCTGAGCACAATTATAATACAG 360
Db 301 CCAACCTCAACTAAGGACAGAGAGACTATAAAATTTGCTGAGCACAATTATAATACAG 360
Qy 361 AGATCTTGAAAGTATTGATAATGAGTGGAGAGAGACTCAATCATGCCACGGAGGTGT 420
Db 361 AGATCTTGAAAGTATTGATAATGAGTGGAGAGAGACTCAATCATGCCACGGAGGTGT 420
Qy 421 GTATAGATGTGGGAAGGAGTTTGGAGTCGCGACAAACACCTCTTTTAAACCTCCATGTG 480
Db 421 GTATAGATGTGGGAAGGAGTTTGGAGTCGCGACAAACACCTCTTTTAAACCTCCATGTG 480
Qy 481 TGTCCGCTACAGATGTGGGGTTGTGCAATAGTAGGGGCTGCAGTGCATGAACACCA 540
Db 481 TGTCCGCTACAGATGTGGGGTTGTGCAATAGTAGGGGCTGCAGTGCATGAACACCA 540
Qy 541 GCACGAGCTACTCTCAGCAAGACGTTATTGAAATTTACAGTGCCTCTCTCTCAAGGCCCA 600
Db 541 GCACGAGCTACTCTCAGCAAGACGTTATTGAAATTTACAGTGCCTCTCTCTCAAGGCCCA 600
Qy 601 AACAGTAAACATCAGTTTGGCAATCAGACTTCTGCGCATGATGATCTTAAACTGGATG 660
Db 601 AACAGTAAACATCAGTTTGGCAATCAGACTTCTGCGCATGATGATCTTAAACTGGATG 660

RESULT 8
US-10-346-802-1
; Sequence 1, Application US/10346802
; Publication No. US20030166873A1

661 TTTACAGACAAGTTCATTTCATTTAGAGCTTCCCTGCCAGCAACACTACCAGTGTG 720
Db TTTACAGACAAGTTCATTTCATTTAGAGCTTCCCTGCCAGCAACACTACCAGTGTG 720
Qy 721 AGGAGGCAACAAGACCTGCCCACTACATGTGGAATAATCACTCTGCAAGATGCC 780
Db 721 AGGAGGCAACAAGACCTGCCCACTACATGTGGAATAATCACTCTGCAAGATGCC 780
Qy 781 TGGCTCAGGAAGATTTTATGTTTCTCGGATGCTGGAGATGCTCAACAGATGGATTC 840
Db 781 TGGCTCAGGAAGATTTTATGTTTCTCGGATGCTGGAGATGCTCAACAGATGGATTC 840
Qy 841 ATGACATCTGTGGACCAACAAGAGCTGGATGAAGAGACCTGTCACTGTCTGCGAG 900
Db 841 ATGACATCTGTGGACCAACAAGAGCTGGATGAAGAGACCTGTCACTGTCTGCGAG 900
Qy 901 CGGGCTTCGGCTGCGAGCTGTGGAACCAACAAGAACTAGACAGAACTCATGCGAGT 960
Db 901 CGGGCTTCGGCTGCGAGCTGTGGAACCAACAAGAACTAGACAGAACTCATGCGAGT 960
Qy 961 GTGCTGTAAAAACAACACTCTTCCCGAGCAATGTGGGCGCAACCGAGAAATTTGATGAA 1020
Db 961 GTGCTGTAAAAACAACACTCTTCCCGAGCAATGTGGGCGCAACCGAGAAATTTGATGAA 1020
Qy 1021 ACACATGCCAGTGTGTATGTAAAAAGAACTGCGCCAGAAATCAACCCCTAAATCCTGAA 1080
Db 1021 ACACATGCCAGTGTGTATGTAAAAAGAACTGCGCCAGAAATCAACCCCTAAATCCTGAA 1080
Qy 1081 AATGTCCTGTGAATGTACAGAAAGTCCAGAAATGCTTGTAAAAAGAAAGATGCC 1140
Db 1081 AATGTCCTGTGAATGTACAGAAAGTCCAGAAATGCTTGTAAAAAGAAAGATGCC 1140
Qy 1141 ACCACCAACATGAGTGTACAGAGCGCCATGTACGAAACCGCAGAGGCTTGTGAGC 1200
Db 1141 ACCACCAACATGAGTGTACAGAGCGCCATGTACGAAACCGCAGAGGCTTGTGAGC 1200
Qy 1201 CAGGATTTTCATATAGTGAAGAGTGTGCTGTTGCTGCTTCCCTTATTTGGCAAGACAC 1260
Db 1201 CAGGATTTTCATATAGTGAAGAGTGTGCTGTTGCTGCTTCCCTTATTTGGCAAGACAC 1260
Qy 1261 AAATGAGCTAAGATTTACTGTTTCCAGTTCATCGATTTTCTATTATGGAATACTGTGT 1320
Db 1261 AAATGAGCTAAGATTTACTGTTTCCAGTTCATCGATTTTCTATTATGGAATACTGTGT 1320
Qy 1321 TGCCACAGTAGAAGTGTGTGAACAGAGAGACCTTGTGGGTCAATGCTAACAAAGACA 1380
Db 1321 TGCCACAGTAGAAGTGTGTGAACAGAGAGACCTTGTGGGTCAATGCTAACAAAGACA 1380
Qy 1381 AAAGTCTCTTCTGTAACCATGTTGTAACCTTTACAGAAATGAGCTGGAGCTCATCTG 1440
Db 1381 AAAGTCTCTTCTGTAACCATGTTGTAACCTTTACAGAAATGAGCTGGAGCTCATCTG 1440
Qy 1441 CAAAAGGCTCTTGTAAAGACTGGTTTCTGCCAATGACCAACAGCCAGATTTTCTC 1500
Db 1441 CAAAAGGCTCTTGTAAAGACTGGTTTCTGCCAATGACCAACAGCCAGATTTTCTC 1500
Qy 1501 TTGGAATTTCTTTAAAGAAATGACTATAATAATTTTCCACTAAAATATTGTTCTGC 1560
Db 1501 TTGGAATTTCTTTAAAGAAATGACTATAATAATTTTCCACTAAAATATTGTTCTGC 1560
Qy 1561 ATTCAATTTTATAGCAACAATTTGTAATAACTCAGTGTGATCAATATTTTATATCAT 1620
Db 1561 ATTCAATTTTATAGCAACAATTTGTAATAACTCAGTGTGATCAATATTTTATATCAT 1620
Qy 1621 GCAAAATATGTTTAAATTAATGAAATTTGATTTTATAAAAAAATAAAAAA 1674
Db 1621 GCAAAATATGTTTAAATTAATGAAATTTGATTTTATAAAAAAATAAAAAA 1674

Db	951	AGCAGCGAAACAAGACCTGCCCCACCAATTTACATGTGGAAATATCATCTGCAGATGCC	892
Qy	781	TGGCTCAGGAAGATTTATGTTTTCTCGGATGCTGGAGATGACTCAACAGATGGATTCC	840
Db	891	TGGCTCAGGAGATTTATGTTTTCTCGGATGCTGGAGATGACTCAACAGATGGATTCC	832
Qy	841	ATGACATCTGTGGAACAAACAAGAGCTGGATGAAGAGACCTGTCAGTGTGTCTGCAGAG	900
Db	831	ATGACATCTGTGACCAAAACAAGAGCTGGATGAAGAGACCTGTCAGTGTGTCTGCAGAG	772
Qy	901	CGGGGCTTCGGCTGCCAGCTGTGGACCCACAAAGAACTAGACAGAAACTCATGCCAGT	960
Db	771	CGGGGCTTCGGCTGCCAGCTGTGGACCCACAAAGAACTAGACAGAAACTCATGCCAGT	712
Qy	961	GTGCTCTGTA AAAA CAAACTCTTCCCAAGCCAAATGTGGGGCCAAACCGAGAAATTTGATGAAA	1020
Db	711	GTGCTCTGTA AAAA CAAACTCTTCCCAAGCCAAATGTGGGGCCAAACCGAGAAATTTGATGAAA	652
Qy	1021	ACACATGCCAGTGTGATGTAAAAGAACCTGCCCCACAGAAATCAACCCCTAAATCTCGGAA	1080
Db	651	ACACATGCCAGTGTGATGTAAAAGAACCTGCCCCACAGAAATCAACCCCTAAATCTCGGAA	592
Qy	1081	AATGTGCTGTGTAATGTACAGAAAGTCCACAGAAATGCTTGTTTAAAAGGAAAGAGTTCC	1140
Db	591	AATGTGCTGTGTAATGTACAGAAAGTCCACAGAAATGCTTGTTTAAAAGGAAAGAGTTCC	532
Qy	1141	ACCACAAACATGCAGCTGTTACAGACGGCCATGTACGAACCGCCAGAAAGCTTGTGAGC	1200
Db	531	ACCACAAACATGCAGCTGTTACAGACGGCCATGTACGAACCGCCAGAAAGCTTGTGAGC	472
Qy	1201	CAGGATTTTCATATAGTAGAGAGTGTGCTGCTGCCCTTCATATGCGAAGACCCAC	1260
Db	471	CAGGATTTTCATATAGTAGAGAGTGTGCTGCTGCCCTTCATATGCGAAGACCCAC	412
Qy	1261	AAATGAGCTAAGANTGTACTGTTTTCCAGTTTCATCGATTTTCTATTATGAAAACCTGTGT	1320
Db	411	AAATGAGCTAAGANTGTACTGTTTTCCAGTTTCATCGATTTTCTATTATGAAAACCTGTGT	352
Qy	1321	TGCCACAGTAGAACTGTCTGTGAACAGAGAGACCTTTGTGGGTCCATGCTTAACAAGACA	1380
Db	351	TGCCACAGTAGAACTGTCTGTGAACAGAGAGACCTTTGTGGGTCCATGCTTAACAAGACA	292
Qy	1381	AAAGTCTGTCTTTCCCTGAAACCATGTGGATAACTTTACAGAAATGGAGCTGAGCTCATCTG	1440
Db	291	AAAGTCTGTCTTTCCCTGAAACCATGTGGATAACTTTACAGAAATGGAGCTGAGCTCATCTG	232
Qy	1441	CAAAAGGCCCTCTTGTAAAAGACTGGTTTTCTGCCAATGACCAACAGCCAAAGATTTTCCCTC	1500
Db	231	CAAAAGGCCCTCTTGTAAAAGACTGGTTTTCTGCCAATGACCAACAGCCAAAGATTTTCCCTC	172
Qy	1501	TTGTGATTTCTTTAAAAGAAATGACTATATAATTTATTTCCACTTAAAAATAATGCTTCTGC	1560
Db	171	TTGTGATTTCTTTAAAAGAAATGACTATATAATTTATTTCCACTTAAAAATAATGCTTCTGC	112
Qy	1561	ATTTCATTTTATAGCAACCAACATTCGTTAAAGCTCACTGTGATCAATATTTTTCATATCAT	1620
Db	111	ATTTCATTTTATAGCAACCAACATTCGTTAAAGCTCACTGTGATCAATATTTTTCATATCAT	52
Qy	1621	GCAAAATATGTTTAAAATAAAAATGAAATTTGTAATTTTATAAAAAA AAAAAA 1671	
Db	51	GCAAAATATGTTTAAAATAAAAATGAAATTTGTAATTTTATAAAAAA AAAAAA AAAAAA	

RESULT 11
US-09-795-006A-21
; Sequence 21, Application US/09795006A
; Patent No. US20020151680A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo et al
; TITLE OF INVENTION: MATERIALS AND METHODS INVOLVING HYBRID VASCULAR
; TITLE OF INVENTION: ENDOTHELIAL GROWTH FACTOR DNAs AND PROTEINS
; FILE REFERENCE: 28967/35977B

CURRENT APPLICATION NUMBER: US/09/795,006A
 CURRENT FILING DATE: 2001-02-26
 PRIOR APPLICATION NUMBER: US 60/205,331
 PRIOR FILING DATE: 2000-05-18
 PRIOR APPLICATION NUMBER: US 60/185,205
 PRIOR FILING DATE: 2000-02-25
 NUMBER OF SEQ ID NOS: 175
 SOFTWARE: PatentIn ver. 2.0
 SEQ ID NO 21
 LENGTH: 1997
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (352)..(1608)
 US-09-795-006A-21

Query Match 98.7%; Score 1651.8; DB 9; Length 1997;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	GTCCCTCCACCATGCATCGTGGCTCTCTCTGTGGGTGTCTCTGCTCGCGCTG	60
DB	341	GTCCCTCCACCATGCATCGTGGCTCTCTCTGTGGGTGTCTCTGCTCGCGCTG	400
QY	61	CGCTGCTCCCGGTCCTCGGAGCGCCGCGCGCGCGCTTCGAGTCCGGACTCG	120
DB	401	CGCTGCTCCCGGTCCTCGGAGCGCCGCGCGCGCGCTTCGAGTCCGGACTCG	450
QY	121	ACCTCTCGGAGCGGAGCCGAGCCGCGCGGAGGCGCACTGTTATGCAAGCAAGATCTGG	180
DB	461	ACCTCTCGGAGCGGAGCCGAGCCGCGCGGAGGCGCACTGTTATGCAAGCAAGATCTGG	520
QY	181	AGGAGCAGTACGCTGTGTGTCAGTGTAGTGAATCATGCTGACTCTCCAGAAAT	240
DB	521	AGGAGCAGTACGCTGTGTGTCAGTGTAGTGAATCATGCTGACTCTCCAGAAAT	580
QY	241	ATTGGAAAATGTACAAGTGTGCTAGTAAAGGAGGCTGGCAACATAACAGAGAACAGG	300
DB	581	ATTGGAAAATGTACAAGTGTGCTAGTAAAGGAGGCTGGCAACATAACAGAGAACAGG	640
QY	301	CCACCTCACTCAAGGACAGAGAGACTATTAATTTGTCGAGCAGCATATTAATACAG	360
DB	641	CCACCTCACTCAAGGACAGAGAGACTATTAATTTGTCGAGCAGCATATTAATACAG	700
QY	361	AGATCTTGAAGAATGATTAATAGTGGAGAAAGACTCAATGATGATGCCACGGAGGTGT	420
DB	701	AGATCTTGAAGAATGATTAATAGTGGAGAAAGACTCAATGATGATGCCACGGAGGTGT	760
QY	421	GTATAGATGGGGAAGAGTTTGGAGTCCGCAACAAACACCTCTTTAAACCTCCATGTG	480
DB	761	GTATAGATGGGGAAGAGTTTGGAGTCCGCAACAAACACCTCTTTAAACCTCCATGTG	820
QY	481	TGTCGGTCTACAGATGTGGGGTTGCTGCAATAGTGGGGGCTGCATGATCAACACCA	540
DB	821	TGTCGGTCTACAGATGTGGGGTTGCTGCAATAGTGGGGGCTGCATGATCAACACCA	880
QY	541	GCACGAGTACTCTCAGCAAGACGTTATTTGAAATTTACAGTGCCTCTCTCTCAAGGCCCA	600
DB	881	GCACGAGTACTCTCAGCAAGACGTTATTTGAAATTTACAGTGCCTCTCTCTCAAGGCCCA	940
QY	601	AACCAAGTAACATCAGTTTTCCTCAATCAGCTTCCCTGCGCATGCTCTAACTCGATG	660
DB	941	AACCAAGTAACATCAGTTTTCCTCAATCAGCTTCCCTGCGCATGCTCTAACTCGATG	1000
QY	661	TTTACAGCAAGTTTCATTCCATTTATAGAGTTTCCCTGCCAGCAACACTACCACTGTC	720
DB	1001	TTTACAGCAAGTTTCATTCCATTTATAGAGTTTCCCTGCCAGCAACACTACCACTGTC	1060
QY	721	AGGAGCGGAACAAGACCTGCCCAACAAATACATGTGGAATATCATCTGCAGATGCC	780
DB	1061	AGGAGCGGAACAAGACCTGCCCAACAAATACATGTGGAATATCATCTGCAGATGCC	1120

RESULT 12
 US-09-375-248-3
 ; Sequence 3, Application US/09375248
 ; Publication No. US20030026759A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ferrell, Robert E.
 ; APPLICANT: Alitalo, Kari
 ; APPLICANT: Finesold, David N.
 ; APPLICANT: Karkkainen, Marika
 ; TITLE OF INVENTION: SCREENING AND THERAPY FOR LYPHATIC DISORDERS INVOLVING
 ; THE FLTA RECEPTOR TYROSINE KINASE (VEGFR-3)
 ; FILE REFERENCE: 28967/35255A

QY	781	TGCTCAGAGAGATTTTATGTGTTTCTCGATGCTGGAGATGATCTCAACAGATGATTC	840
DB	1121	TGCTCAGAGAGATTTTATGTGTTTCTCGATGCTGGAGATGATCTCAACAGATGATTC	1180
QY	841	ATGACATCTGTGACCAACAAAGAGCTGGATGAAGAGACCTGTGAGTGTGTCTCAGAG	900
DB	1181	ATGACATCTGTGACCAACAAAGAGCTGGATGAAGAGACCTGTGAGTGTGTCTCAGAG	1240
QY	901	CGGGCTTGGCTCGCCTGCTGAGCCCCCAAGAACTAGACAGAACTCATGCCAGT	960
DB	1241	CGGGCTTGGCTCGCCTGCTGAGCCCCCAAGAACTAGACAGAACTCATGCCAGT	1300
QY	961	GTGCTGTGTAACAAACAACTCTTCCCGAGCCATGTGGGCGCAACCGAGAAATTTGATGAA	1020
DB	1301	GTGCTGTGTAACAAACAACTCTTCCCGAGCCATGTGGGCGCAACCGAGAAATTTGATGAA	1360
QY	1021	ACACATGCCAGTGTGTATGTAAGAACTGCCCCAGAAATCAACCCCTAAATCCTGGAA	1080
DB	1361	ACACATGCCAGTGTGTATGTAAGAACTGCCCCAGAAATCAACCCCTAAATCCTGGAA	1420
QY	1081	ATGTGCTGTGTAAGTAACAGAACTCCACAGAAATGCTTGTAAAGGAAGAGTTTC	1140
DB	1421	ATGTGCTGTGTAAGTAACAGAACTCCACAGAAATGCTTGTAAAGGAAGAGTTTC	1480
QY	1141	ACCACCAACATGACAGCTGTTACAGCGCCCATGTACGAAACCGCCAGAGGCTTGTGAGC	1200
DB	1481	ACCACCAACATGACAGCTGTTACAGCGCCCATGTACGAAACCGCCAGAGGCTTGTGAGC	1540
QY	1201	CAGGATTTTCTATAGTGAAGAGTGTGTCCTTTCATATTTGGGCAAGACAC	1260
DB	1541	CAGGATTTTCTATAGTGAAGAGTGTGTCCTTTCATATTTGGGCAAGACAC	1600
QY	1261	AAATCAGCTAAGATGTACTGTTTCCAGTTTCATGATTTTCTATATGGAAGAACTGTGT	1320
DB	1601	AAATCAGCTAAGATGTACTGTTTCCAGTTTCATGATTTTCTATATGGAAGAACTGTGT	1660
QY	1321	TGCCACAGTACAGTGTCTGTGAACAGAGAGACCTTGTGGTGCATGCTCAACAAAGACA	1380
DB	1661	TGCCACAGTACAGTGTCTGTGAACAGAGAGACCTTGTGGTGCATGCTCAACAAAGACA	1720
QY	1381	AAAGTCTCTCTTCTCTGAAACCATGTGGATTAATTTACAGAAATGAGCTGGAGCTCATCTG	1440
DB	1721	AAAGTCTCTCTTCTCTGAAACCATGTGGATTAATTTACAGAAATGAGCTGGAGCTCATCTG	1780
QY	1441	CAAAAGGCTCTTGTAAAGACTGTTTTCTGCCAATGACCAACAGCCAGCAAGTTTCTC	1500
DB	1781	CAAAAGGCTCTTGTAAAGACTGTTTTCTGCCAATGACCAACAGCCAGCAAGTTTCTC	1840
QY	1501	TTGTGATTTCTTAAAGAAATGACTATATATTTTCCACTAAATATTTGTTCTGC	1560
DB	1841	TTGTGATTTCTTAAAGAAATGACTATATATTTTCCACTAAATATTTGTTCTGC	1900
QY	1561	ATTCATTTTATAGCAACAAATTTGTAAGAACTCACTGTGATCAATATTTTATATCAT	1620
DB	1901	ATTCATTTTATAGCAACAAATTTGTAAGAACTCACTGTGATCAATATTTTATATCAT	1960
QY	1621	GCAAAATATGTTTAAATATAAATGAAATTTGATTT	1655
DB	1961	GCAAAATATGTTTAAATATAAATGAAATTTGATTT	1995

Db	1961	GCAAAATATGTTTAAATAAATAAATGAAATGTATT	1995
Db	1001	TTTACAGACAAGTTCATTCCATTATTAGACGTTCCCTGCAGCAACACTACCAAGTTC	1060
Qy	721	AGGACGCAACAAAGACTGCTCCCAACAAATTACATGTGGAAATAATACATCTGCAGATGCC	780
Db	1061	AGGACGCAACAAAGACTGCTCCCAACAAATTACATGTGGAAATAATACATCTGCAGATGCC	1120
Qy	781	TGGCTCAGGAAGATTTTATGTTTTCCTCGATGCTGGAGATGACTCAACAGATGGATTC	840
Db	1121	TGGCTCAGGAAGATTTTATGTTTTCCTCGATGCTGGAGATGACTCAACAGATGGATTC	1180
Qy	841	ATGACATCTGTGACCAACAAAGACTGCTGGAGATGACTCAACAGATGGATTC	900
Db	1181	ATGACATCTGTGACCAACAAAGACTGCTGGAGATGACTCAACAGATGGATTC	1240
Qy	901	CGGGCTTCGGCTGCGCAGCTGTGGACCCCAACAAAGACTTAGACAGAACTCATGCCAGT	960
Db	1241	CGGGCTTCGGCTGCGCAGCTGTGGACCCCAACAAAGACTTAGACAGAACTCATGCCAGT	1300
Qy	961	GTGCTGTAAACAAACTCTTCCCAAGCAATGTGGGCAACCGAGAAATTTGATGAA	1020
Db	1301	GTGCTGTAAACAAACTCTTCCCAAGCAATGTGGGCAACCGAGAAATTTGATGAA	1360
Qy	1021	ACATATGCCAGTGTATGTAAAGAACTCTGCCCCAGAAATCAACCCCTAAATCTCTGAA	1080
Db	1361	ACATATGCCAGTGTATGTAAAGAACTCTGCCCCAGAAATCAACCCCTAAATCTCTGAA	1420
Qy	1081	AATGCTCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTTAAAGAAAGAAAGTTC	1140
Db	1421	AATGCTCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTTAAAGAAAGAAAGTTC	1480
Qy	1141	ACCACCAACATGACGCTGTACAGAGCGCCATGTACGAAACCGCCAGAGAGCTTGTGAGC	1200
Db	1481	ACCACCAACATGACGCTGTACAGAGCGCCATGTACGAAACCGCCAGAGAGCTTGTGAGC	1540
Qy	1201	CAGGATTTTATATAGTGAAGAGTGTGCTGTGTGCTCCCTTATATTCGCAAGAGCTG	1260
Db	1541	CAGGATTTTATATAGTGAAGAGTGTGCTGTGTGCTCCCTTATATTCGCAAGAGCTG	1600
Qy	1261	AAATGAGCTAAAGTGTACTGTTTCCAGTTCATCGATTTTCTATTTATGGAAGACTGT	1320
Db	1601	AAATGAGCTAAAGTGTACTGTTTCCAGTTCATCGATTTTCTATTTATGGAAGACTGT	1660
Qy	1321	TGCCAGTAGAAGTCTGCTGTGAACAGAGAGACCTTGTGGGTCCATGTCTAAAGAGACA	1380
Db	1661	TGCCAGTAGAAGTCTGCTGTGAACAGAGAGACCTTGTGGGTCCATGTCTAAAGAGACA	1720
Qy	1381	AAAGTCTGCTTTTCTGAAACCATGTGGATTAATTACAGAAATGGAGCTCATCTG	1440
Db	1721	AAAGTCTGCTTTTCTGAAACCATGTGGATTAATTACAGAAATGGAGCTCATCTG	1780
Qy	1441	CAAAAGGCTCTGTGAAAGACTGTTTCTGCAATGACCAACAGCCAGATTTTCTCTC	1500
Db	1781	CAAAAGGCTCTGTGAAAGACTGTTTCTGCAATGACCAACAGCCAGATTTTCTCTC	1840
Qy	1501	TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTTCCACTAAATAATTTTCTG	1560
Db	1841	TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTTCCACTAAATAATTTTCTG	1900
Qy	1561	ATTCAATTTTATAGCAACAAATTTGGTAAACTCAGTGTGATCAATTTTATATCAT	1620
Db	1901	ATTCAATTTTATAGCAACAAATTTGGTAAACTCAGTGTGATCAATTTTATATCAT	1960
Qy	1621	GCAAAATATGTTTAAATAAATAAATGAAATGTATT	1655
Db	1961	GCAAAATATGTTTAAATAAATAAATGAAATGTATT	1995

RESULT 15
 US-10-081-126-3
 ; Sequence 3, Application US/10081126
 ; Publication No. US20030180294A1
 ; GENERAL INFORMATION:

Db	1961	GCAAAATATGTTTAAATAAATAAATGAAATGTATT	1995
Db	1001	TTTACAGACAAGTTCATTCCATTATTAGACGTTCCCTGCAGCAACACTACCAAGTTC	1060
Qy	721	AGGACGCAACAAAGACTGCTCCCAACAAATTACATGTGGAAATAATACATCTGCAGATGCC	780
Db	1061	AGGACGCAACAAAGACTGCTCCCAACAAATTACATGTGGAAATAATACATCTGCAGATGCC	1120
Qy	781	TGGCTCAGGAAGATTTTATGTTTTCCTCGATGCTGGAGATGACTCAACAGATGGATTC	840
Db	1121	TGGCTCAGGAAGATTTTATGTTTTCCTCGATGCTGGAGATGACTCAACAGATGGATTC	1180
Qy	841	ATGACATCTGTGACCAACAAAGACTGCTGGAGATGACTCAACAGATGGATTC	900
Db	1181	ATGACATCTGTGACCAACAAAGACTGCTGGAGATGACTCAACAGATGGATTC	1240
Qy	901	CGGGCTTCGGCTGCGCAGCTGTGGACCCCAACAAAGACTTAGACAGAACTCATGCCAGT	960
Db	1241	CGGGCTTCGGCTGCGCAGCTGTGGACCCCAACAAAGACTTAGACAGAACTCATGCCAGT	1300
Qy	961	GTGCTGTAAACAAACTCTTCCCAAGCAATGTGGGCAACCGAGAAATTTGATGAA	1020
Db	1301	GTGCTGTAAACAAACTCTTCCCAAGCAATGTGGGCAACCGAGAAATTTGATGAA	1360
Qy	1021	ACATATGCCAGTGTATGTAAAGAACTCTGCCCCAGAAATCAACCCCTAAATCTCTGAA	1080
Db	1361	ACATATGCCAGTGTATGTAAAGAACTCTGCCCCAGAAATCAACCCCTAAATCTCTGAA	1420
Qy	1081	AATGCTCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTTAAAGAAAGAAAGTTC	1140
Db	1421	AATGCTCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTTAAAGAAAGAAAGTTC	1480
Qy	1141	ACCACCAACATGACGCTGTACAGAGCGCCATGTACGAAACCGCCAGAGAGCTTGTGAGC	1200
Db	1481	ACCACCAACATGACGCTGTACAGAGCGCCATGTACGAAACCGCCAGAGAGCTTGTGAGC	1540
Qy	1201	CAGGATTTTATATAGTGAAGAGTGTGCTGTGTGCTCCCTTATATTCGCAAGAGCTG	1260
Db	1541	CAGGATTTTATATAGTGAAGAGTGTGCTGTGTGCTCCCTTATATTCGCAAGAGCTG	1600
Qy	1261	AAATGAGCTAAAGTGTACTGTTTCCAGTTCATCGATTTTCTATTTATGGAAGACTGT	1320
Db	1601	AAATGAGCTAAAGTGTACTGTTTCCAGTTCATCGATTTTCTATTTATGGAAGACTGT	1660
Qy	1321	TGCCAGTAGAAGTCTGCTGTGAACAGAGAGACCTTGTGGGTCCATGTCTAAAGAGACA	1380
Db	1661	TGCCAGTAGAAGTCTGCTGTGAACAGAGAGACCTTGTGGGTCCATGTCTAAAGAGACA	1720
Qy	1381	AAAGTCTGCTTTTCTGAAACCATGTGGATTAATTACAGAAATGGAGCTCATCTG	1440
Db	1721	AAAGTCTGCTTTTCTGAAACCATGTGGATTAATTACAGAAATGGAGCTCATCTG	1780
Qy	1441	CAAAAGGCTCTGTGAAAGACTGTTTCTGCAATGACCAACAGCCAGATTTTCTCTC	1500
Db	1781	CAAAAGGCTCTGTGAAAGACTGTTTCTGCAATGACCAACAGCCAGATTTTCTCTC	1840
Qy	1501	TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTTCCACTAAATAATTTTCTG	1560
Db	1841	TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTTCCACTAAATAATTTTCTG	1900
Qy	1561	ATTCAATTTTATAGCAACAAATTTGGTAAACTCAGTGTGATCAATTTTATATCAT	1620
Db	1901	ATTCAATTTTATAGCAACAAATTTGGTAAACTCAGTGTGATCAATTTTATATCAT	1960
Qy	1621	GCAAAATATGTTTAAATAAATAAATGAAATGTATT	1655
Db	1961	GCAAAATATGTTTAAATAAATAAATGAAATGTATT	1995

RESULT 15
 US-10-081-126-3
 ; Sequence 3, Application US/10081126
 ; Publication No. US20030180294A1
 ; GENERAL INFORMATION:

APPLICANT: De Vries, Gerald W.
TITLE OF INVENTION: Methods of Extending Corneal Graft
FILE REFERENCE: P-AR 4951
CURRENT APPLICATION NUMBER: US/10/081.126
CURRENT FILING DATE: 2002-02-22
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 2015
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (372)...(1628)
US-10-081-126-3

Query Match 98.7%; Score 1651.8; DB 14; Length 2015;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
1 GTCCTTCACCATCAGCTCGCTGGGCTTCTCTCTGTGGGTGTTCTCTGCTCGCGCTG 60
Db GTCTTTCACCATCAGCTCGCTGGGCTTCTCTCTGTGGGTGTTCTCTGCTCGCGCTG 420
61 CGTGTCTCCGGTCTCTCGAGGCGCCGCGCGCGCGCGCGCGCTTCGAGTCCGAGCTCG 120
Db CGTGTCTCCGGTCTCTCGAGGCGCCGCGCGCGCGCGCGCGCTTCGAGTCCGAGCTCG 480
121 ACCTCTCGGACGGGACCGGACCGGCGGCGGCGGCGGCTTATGCGAAGATCTGG 180
Db ACCTCTCGGACGGGACCGGACCGGCGGCGGCGGCGGCTTATGCGAAGATCTGG 540
181 AGGAGCAGTTACGGTCTGTGTCCAGTGTAGATGAATCATGACTGTACTCTACCCAGAAT 240
Db AGGAGCAGTTACGGTCTGTGTCCAGTGTAGATGAATCATGACTGTACTCTACCCAGAAT 600
241 ATTGGAAAATGTACAAGTGTACGCTAAGGAAAGAGGCTGGCAACATAACAGAGAACAG 300
Db ATTGGAAAATGTACAAGTGTACGCTAAGGAAAGAGGCTGGCAACATAACAGAGAACAG 660
301 CCAACCTCACTCAAGCAGACAGAGACTATAAATTTGCTGCGAGCAGCATTAATACAG 360
Db CCAACCTCACTCAAGCAGACAGAGACTATAAATTTGCTGCGAGCAGCATTAATACAG 720
361 AGATCTTGAAGATTAATGATAGTGGAGAAAGACTCAATGCATGCCACGGGAGGTGT 420
Db AGATCTTGAAGATTAATGATAGTGGAGAAAGACTCAATGCATGCCACGGGAGGTGT 780
421 GTATAGATGGGGAAGGATTTGGAGTCGGACAAACACCTTCTTTAAACCTCCATGTG 480
Db GTATAGATGGGGAAGGATTTGGAGTCGGACAAACACCTTCTTTAAACCTCCATGTG 840
481 TGTCGGCTCAGATGTGGGGTGTCTGCAATAGTAGAGGGGCTGCAGTGCATGAACACA 540
Db TGTCGGCTCAGATGTGGGGTGTCTGCAATAGTAGAGGGGCTGCAGTGCATGAACACA 900
541 GCAGAGCTACCTCAGCAGACAGTTATTGAAATTAACAGTCTCTCTCTCAAGGCCCA 600
Db GCAGAGCTACCTCAGCAGACAGTTATTGAAATTAACAGTCTCTCTCTCAAGGCCCA 960
601 AACCAAGTAAACATCAGTTTGGCAATCACAATCTCTGCGGATCGATGCTTAACTGGATG 660
Db AACCAAGTAAACATCAGTTTGGCAATCACAATCTCTGCGGATCGATGCTTAACTGGATG 1020
661 TTACAGACAGTTTCAATTCATTAATAGAGCTTCCCTGCGAGCAACACTACACAGTGTG 720
Db TTACAGACAGTTTCAATTCATTAATAGAGCTTCCCTGCGAGCAACACTACACAGTGTG 1080
721 AGGAGGGAACAGAGCTGCGGACCAATTAACATGAGATTAATCACTGCGAGTGC 780
Db AGGAGGGAACAGAGCTGCGGACCAATTAACATGAGATTAATCACTGCGAGTGC 1140

QY 781 TGGCTCAGGAAGATTTTATGTTTTCTCGGATGCTGGAGATGACTCAACAGATGATTC 840
Db TGGCTCAGGAAGATTTTATGTTTTCTCGGATGCTGGAGATGACTCAACAGATGATTC 1200
QY 841 ATGACATCTGTGGACCAACAAAGAGCTGGATGAAGAGACCTGTCAAGTGTGCTCAGAG 900
Db ATGACATCTGTGGACCAACAAAGAGCTGGATGAAGAGACCTGTCAAGTGTGCTCAGAG 1260
QY 901 CGGGCTTGGCTCGGCTCGGACCCCAAGAACTAGACAGAACTCATCCAGT 960
Db CGGGCTTGGCTCGGCTCGGACCCCAAGAACTAGACAGAACTCATCCAGT 1320
QY 961 GTGCTGTAAACAAACCTCTTCCCGACGATGTGGGCGCAACCGAGAACTTGATGAA 1020
Db GTGCTGTAAACAAACCTCTTCCCGACGATGTGGGCGCAACCGAGAACTTGATGAA 1380
QY 1021 ACACATGCCAGTGTGTATGTAAGAACTCTGCCAGAAATCAACCCCTAAATCTCGAA 1080
Db ACACATGCCAGTGTGTATGTAAGAACTCTGCCAGAAATCAACCCCTAAATCTCGAA 1440
QY 1081 ATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTTAAAGGAAGAAAGTTC 1140
Db ATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTTAAAGGAAGAAAGTTC 1500
QY 1441 AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTTAAAGGAAGAAAGTTC 1200
Db AATGTGCTGTGAATGTACAGAAAGTCCACAGAAATGCTTTTAAAGGAAGAAAGTTC 1560
QY 1141 ACCACCAACATGCACTGTTCACAGCGCCATGTACGAAACCGCCAGAAAGCTTGTGAGC 1260
Db ACCACCAACATGCACTGTTCACAGCGCCATGTACGAAACCGCCAGAAAGCTTGTGAGC 1620
QY 1201 CAGGATTTTCAATAGTGAAGAGTGTGCTGTGCTCCCTTCATATTTGGCAAGACAC 1320
Db CAGGATTTTCAATAGTGAAGAGTGTGCTGTGCTCCCTTCATATTTGGCAAGACAC 1680
QY 1261 AAATGAGCTAAGATTTGACTGTTTCCAGTTTCATCGATTTTCTATTTATGGAAGTGT 1380
Db AAATGAGCTAAGATTTGACTGTTTCCAGTTTCATCGATTTTCTATTTATGGAAGTGT 1740
QY 1321 TGCCACAGTGAACCTGTCTGTGAACAGAGAGACCTTTGTGGGTCCATGCTAACAAAGACA 1440
Db TGCCACAGTGAACCTGTCTGTGAACAGAGAGACCTTTGTGGGTCCATGCTAACAAAGACA 1800
QY 1381 AAAGTCTGTCTTCTGAAACCATGTGGATTAATTTTACAGAAATGAGCTGGAGCTCATCTG 1500
Db AAAGTCTGTCTTCTGAAACCATGTGGATTAATTTTACAGAAATGAGCTGGAGCTCATCTG 1860
QY 1441 CAAAGGGCTCTTGTAAAGACTGGTTTTCTGCCAATGACCAAAACAGCCAGATTTTCTC 1560
Db CAAAGGGCTCTTGTAAAGACTGGTTTTCTGCCAATGACCAAAACAGCCAGATTTTCTC 1920
QY 1501 TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAAAATATTTGTTCTGC 1620
Db TTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAAAATATTTGTTCTGC 1980
QY 1561 ATTCAATTTTATAGCAACAAATTTGGTAAACCTCACTGTGATCAATATTTTATATCAT 1620
Db ATTCAATTTTATAGCAACAAATTTGGTAAACCTCACTGTGATCAATATTTTATATCAT 1980
QY 1621 GCAAAATATGTTTAAATTAATGAAATGAAATGATTT 1655
Db GCAAAATATGTTTAAATTAATGAAATGAAATGATTT 2015

Search completed: February 26, 2004, 01:43:17
Job time : 627 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2004, 16:56:30 ; Search time 3995 Seconds
(without alignments)
12512.964 Million cell updates/sec

Title: US-09-921-143-1
Perfect score: 1674
Sequence: 1 gctcttcacatgcacgcg.....ttataaaaaaaaaaaaaa 1674

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:

1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_hcc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rtd.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1207.6	72.1	2450	11 AK047844	AK047844 Mus muscu
2	839.4	50.1	998	12 BG677890	BG677890 602625584
3	782.4	46.7	831	12 BG680919	BG680919 602628716
4	780.8	46.6	858	14 CA488579	CA488579 AGENCOURT

C	5	730.6	43.6	750	14	CA421508
C	6	729.6	43.6	748	13	BU615444
C	7	702	41.9	808	10	AW195617
C	8	698.6	41.7	717	9	AA923193
C	9	667.2	39.9	728	9	AW070215
C	10	656.2	39.2	662	12	BQ005831
C	11	649	38.8	657	13	BU619991
C	12	616	36.8	637	10	BE218812
C	13	614.6	36.7	622	14	CB854820
C	14	603	36.0	916	13	BU510633
C	15	576	34.4	584	12	BM127186
C	16	575.4	34.4	618	9	AA151613
C	17	570	34.1	578	12	BM054372
C	18	561.2	33.5	575	10	AW341345
C	19	549.8	32.9	983	10	BF669473
C	20	527.8	31.5	564	10	BE670542
C	21	523.4	31.3	525	12	BM055226
C	22	523	31.2	535	13	BM113454
C	23	522.8	31.2	537	9	AI827929
C	24	517.2	30.9	826	10	BE958208
C	25	515.8	30.8	519	12	BM127468
C	26	514.8	30.8	570	9	AI580094
C	27	508.2	30.4	535	9	AI459340
C	28	507.4	30.3	509	9	AA425486
C	29	505	30.2	631	10	BF240811
C	30	500.8	29.9	659	14	CF109307
C	31	494	29.5	502	10	AW291800
C	32	489.8	29.3	593	9	AI243720
C	33	485.6	29.0	562	14	N31713
C	34	483.4	28.9	509	12	BM310538
C	35	477.4	28.5	479	12	BM310845
C	36	472.8	28.2	630	14	N31720
C	37	471	28.1	703	12	BP110008
C	38	470	28.1	504	10	AW959892
C	39	465.6	27.8	472	10	AW137803
C	40	463.4	27.7	598	10	BE222114
C	41	448.6	26.8	1398	12	BM912752
C	42	437.6	26.1	593	10	BE376968
C	43	432.4	25.8	434	9	AA406492
C	44	422.5	25.2	429	9	AI066424
C	45	413.2	24.7	973	12	BI154009

ALIGNMENTS

RESULT 1
AK047844
LOCUS
DEFINITION
AK047844 2450 bp mRNA linear HTC 20-SEP-2003
Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched
library, clone:Cl30004M09 product:vascular endothelial growth
factor C, full insert sequence.

ACCESSION
AK047844.1 GI:26339001

VERSION
AK047844.1

KEYWORDS
HTC; CAP trapper.

SOURCE
Mus musculus (house mouse)

ORGANISM
Mus musculus

REFERENCE
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
PUBMED 10349636

REFERENCE
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
PUBMED 11042159

REFERENCE
AUTHORS

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Iizawa, M., Ohara, E., Watabiki, M.,
Yoneda, Y., Iehikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
RIKEN integrated sequence analysis (KISA) system--384-format
sequencing pipeline with 384 multipillar sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

REFERENCE
AUTHORS

4 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

JOURNAL
MEDLINE
PUBMED

5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2450)

JOURNAL
REFERENCE
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Atakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kigawa, I., Kasukawa, T.,
Kato, S., Kurihara, C., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koyama, M., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, I., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M., and Hayashizaki, Y.

JOURNAL
REFERENCE
AUTHORS

Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan [E-mail: genome-res@sc.riken.go.jp].
URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222,
Fax: 81-45-503-9216

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>.

FEATURES
source

Location/Qualifiers
1. 2450
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:C130004M09"
/db_xref="MG1:2413670"
/db_xref="taxon:10090"
/clone="C130004M09"
/tissue_type="head"
/clone_libs="RIKEN full-length enriched mouse cDNA library"
/dev_stages="16 days embryo"
175. .1422
/note="unnamed protein product; putative
vascular endothelial growth factor C (MGD|MG1:109124,
GB|NM_009506, evidence: BLASTN, 99%, match=1784)"
/codon_start=1
/protein_id="BAC33172.1"
/db_xref="GI:26339002"
/translation="MHLLCLFLSLACSLIAALIPSPREAPATVAATFESGLGFSEAPDP

ORIGIN

Query Match 72.1%; Score 1207.6; DB 11; Length 2450;
Best Local Similarity 85.6%; Pred. No. 4.7e-188;
Matches 1418; Conservative 0; Mismatches 219; Indels 19; Gaps 6;
QY 1 GTCCCTCCACCATGCACTCGCTGGGCTTCTTCTGTGGCGTGTCTCTGCTCGCGCTG 60
DB 164 GTCCATCCACCATGCACTTGTCTGTCTGTGGCGTGTCTCTGCTCGCGCTG 223
QY 61 CGTGTCTCCCGGCTCCTCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 120
DB 224 CGCTGATCCCACTCGCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 283
QY 121 ACCTCTCGGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 180
DB 284 GTTCTCGGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 343
QY 181 AGGAGCAGTTACGGTCTGTCTCCAGGTAGATGAATCATGACTGTACTTACCAGAAAT 240
DB 344 AGGAGCAGTTGCGGTCTGTCTCCAGGTAGATGAATCATGACTGTCTCTGTTACCCAGCT 403
QY 241 ATTGGAATAATGTAAGTGTACAGTAAAGAAAGAGGCTGGCAACATTAACAGAGAACAGG 300
DB 404 ACTGGAATAATGTAAGTGTACAGTAAAGAAAGAGGCTGGCAACATTAACAGAGAACAGG 451
QY 301 CCAACTCACTCAAGGAGAGAGACTATATAATTTGCTGCAGACATATATATACAG 360
DB 452 CCACCTCAATACCAGGAGGAGAGAGTGTAAATTTGCTGCTGCATATTAACACAG 511
QY 361 AGATCTTGAATAATGTAAGTGTAGTGAAGAAAGAGTCAATGCTGCAGCGGAGGCTG 420
DB 512 AGATCTGTAATAATGTAAGTGTAGTGAAGAAAGAGTCAATGCTGCAGCGGAGGCTG 571
QY 421 GTATAGATGTGGGAGAGGAGTTGGAGTGCAGCAACACCTTCTTTAAACCTCATGTG 480
DB 572 GTATAGATGTGGGAGAGGAGTTGGAGTGCAGCAACACCTTCTTTAAACCTCATGTG 631
QY 481 TGTCCGTCTACAGATGTGGGAGGCTCTGCAATAGTGAAGGCTGCAGTGCATGAACACCA 540
DB 632 TGTCCGTCTACAGATGTGGGAGGCTCTGCAACAGCGAGGCTGCAGTGCATGAACACCA 591
QY 541 GCACGAGTACCTCAGCAAGAGCGTTGTTGAAATTTACAGTGCCTCTCTCAAGGCCCA 600
DB 692 GCACGAGTACCTCAGCAAGAGCGTTGTTGAAATTTACAGTGCCTCTCTCAAGGCCCA 751
QY 601 AACCACTAACATCAGTTTTCGCAATCACCTTCTCGCGATGATGCTAACTGATG 660
DB 752 AACCACTAACATCAGTTTTCGCAATCACCTTCTCGCGATGATGCTAACTGATG 811
QY 661 TTTACAGACAGTTTCAATTCATTTATAGACGTTTCTCTCCAGCAACACTTACACAGTGT 720
DB 812 TTTACAGACAGTTTCAATTCATTTATAGACGTTTCTCTCCAGCAACACTTACACAGTGT 871
QY 721 AGCGAGCGAAGAGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 780
DB 872 AGCGAGCTAACAGAGAGCTGCGCAAACTATGTGTGAATTAACACTTACACTTGCAGATGCC 931
QY 781 TGGCTCAGGAGATTTTATGTTTCTCGGATGCTGGAGATGACTCAACAGATGATGCC 840
DB 932 TGGCTCAGGAGATTTTATCTTTTATCAATGTTGAAGAGTCACTCAACCACTGATGCC 991
QY 841 ATGACATCTGTGGAACAAAGAGGCTGGATGAAGAGAGCTGTGTCAGTGTGTGTGCAAG 900
DB 992 ATGATGTCTGTGGAACCAAGAGGCTGGATGAAGAGAGCTGTGTCAGTGTGTGTGCAAG 1051

QY 901 CGGGCTTGGCTCCAGCTGTGGACCCCAAGAACTAGACAGAACTCATGCCAGT 960
Db 1052 GGGGCTTGGGCTTGTAGTGTGGACCCCAAGAACTAGACAGCTCATGTGAGT 1111
QY 961 GTGTCTGTAAACAACTCTTCCCGAGCGCAATGTGGGCGCAACCGAGAAATTTGATGAA 1020
Db 1112 GTGTCTGTAAACAACTCTTCCCGAGCGCAATGTGGGCGCAACCGAGAAATTTGATGAA 1171
QY 1021 ACATGTCGAGTGTGTATGTAAGAACTGTGCGCCAGAAATCAACCCCTAAATCCTGGAA 1080
Db 1172 ATACATGTGAGTGTGTATGTAAGAACTGTGCGCCAGAAATCAACCCCTAAATCCTGGAA 1231
QY 1081 ATGTGCTGTGTAATGTAGTACAGAACTCCACAGAAATGTGTTAAAGGGAAGAGTTCC 1140
Db 1232 ATGTGCTGTGTAATGTAGTACAGAACTCCACAGAAATGTGTTAAAGGGAAGAGTTCC 1291
QY 1141 ACCACAAACATGTGAGTGTGTACAGAGCGCCATGTACGAACCGCGAGAGCTTTGTGAGC 1200
Db 1292 ACCATCAACATGTGAGTGTGTACAGAGCGCCATGTACGAACCGCGAGAGCTTTGTGATC 1351
QY 1201 CAGGATTTTCATATAGTGAAGAGTGTGCTGTGTCCTTCATATTTGTCCTTCATATTTGGCAAGACCAC 1260
Db 1352 CAGGAGTGTCTTTTGTAGGAGAGTATGCGGTGTGTCCTTCATATTTGTCCTTCATATTTGGCAAGACCAC 1411
QY 1261 AATGAGCTTAAGTGTACTGTTTCCAGTTCATCGATTTTCTATTTGGAATACTGTGT 1320
Db 1412 ATCTGAACATAGATCATACAGTCTTTCAGTCCACATTTTACTCTCTTGAAGAC--TGT 1469
QY 1321 TGCCAGTGTAGTGTGTGTGACAGAGAGACCCCTGTGGGTGCATGCTTAACAAAGACA 1380
Db 1470 TGCCACATGTAGTGTGTGTGACAGAGAGACCCCTGTGGGTGCATGCTTAACAAAGACC 1528
QY 1381 AAGTCTGTCTTCTGAAACATGTGATTAACCTTTACAGAAATGAGCTGGAGCTCATCTG 1440
Db 1529 CAAGTCTGTGTTTGTGAACATGTGATTAACCTTTACAGAAATGAGCTGGAGCTCATCTG 1586
QY 1441 CAAAGGCTCTGTGTAAGAGCTGTTTCTGCCAATGACCAACAGCAAGATTTTCTC 1500
Db 1587 CAAAGGCTCTGTGTAAGAGCTGTTTCTGCCAATGACCAACAGCAAGATTTTCTC 1646
QY 1501 TTGTGATTTCTTAAAGAGTGTATATATTTTTCACCTAAATATTTTCTG 1560
Db 1647 TTGTGATTT-AAAAAAGAGTGTATATATTTTTCACCTAAATATTTTCTG 1705
QY 1561 ATTCAATTTTATAGCAACAACTTGTGTAAGAGCTGATCAATATTTTATATCAT 1620
Db 1706 ATTCAATTTTATAGCAATTAACATTTGTGTAAGAGCTGATCAATATTTTATACAT 1765
QY 1621 CCAAAA--TATGTTTAAATATAAATGAAATTTGTATT 1655
Db 1766 GCAAAACTATGTTTAAATATAAATGAAATTTGTATT 1801

RESULT 2
BG677890
LOCUS 998 bp mRNA linear EST 01-MAY-2001
DEFINITION 602625584F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4750509 5',
mRNA sequence.
ACCESSION BG677890
VERSION BG677890.1 GI:13909287
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 998)
REFERENCE NIH-MGC <http://mgi.nci.nih.gov/>
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.

cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLNL0605 row: g column: 22
High quality sequence stop: 806.
Location/Qualifiers
1..998
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4750509"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Skn4"
/note="Organ: skin; Vector: pCMV-SPORT6, Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

FEATURES
source

ORIGIN

Query Match 50.1%; Score 839.4; DB 12; Length 998;
Best local Similarity 95.4%; Pred. No. 1.1e-127;
Matches 951; Conservative 0; Mismatches 36; Indels 10; Gaps 8;
QY 597 CCCAACCCAGTAACATCAGTTTGGCAATCAGACTTCTCGCGATGATGCTTAAACTG 556
Db 1 CCCAACCCAGTAACATCAGTTTGGCAATCAGACTTCTCGCGATGATGCTTAAACTG 60
QY 657 GATGTTTACAGACAAGTTCATTCCATTATTAGACGTTCCCTGCCAGCAACACTACCACAG 716
Db 61 GATGTTTACAGACAAGTTCATTCCATTATTAGACGTTCCCTGCCAGCAACACTACCACAG 120
QY 717 TGTACGCGAGCAACAGACCTGCCGCCCAATTAATCATGTGGAATTAATCACAATCTGCAGA 776
Db 121 TGTACGCGAGCAACAGACCTGCCGCCCAATTAATCATGTGGAATTAATCACAATCTGCAGA 180
QY 777 TGCTCGCTCAGAGAAATTTATGTTTCTCGGATGCTGGAGATGACATCAACAGATGA 836
Db 181 TGCTCGCTCAGAGAAATTTATGTTTCTCGGATGCTGGAGATGACATCAACAGATGA 240
QY 837 TTCCATGACATCTGTGGACCAACAGAGAGCTGGATGAAGAGACCTGTGAGTGTGTCTGC 896
Db 241 TTCCATGACATCTGTGGACCAACAGAGAGCTGGATGAAGAGACCTGTGAGTGTGTCTGC 300
QY 897 AGAGCGGGCTTGGGCTCCAGCTGTGACCCCAAGAACTAGACAGAACTCATGTC 956
Db 301 AGAGCGGGCTTGGGCTCCAGCTGTGACCCCAAGAACTAGACAGAACTCATGTC 360
QY 957 CAGTGTGTCTGTAAACAACTCTTCCCGAGCAATGTGGGCGCAACCGAGAAATTTGAT 1016
Db 361 CAGTGTGTCTGTAAACAACTCTTCCCGAGCAATGTGGGCGCAACCGAGAAATTTGAT 420
QY 1017 GAAACACATGCCAGTGTGTATGTAAGAACTGTGCGCCAGAAATCAACCCCTAAATCT 1076
Db 421 GAAACACATGCCAGTGTGTATGTAAGAACTGTGCGCCAGAAATCAACCCCTAAATCT 480
QY 1077 GGAATATGCTGTGAATGTACAGAACTGCACAGAAATGCTTGTGTTAA--AGGAAGAA 1135
Db 481 GGAATATGCTGTGAATGTACAGAACTGCACAGAAATGCTTGTGTTAA--AGGAAGAA 540
QY 1136 GTTCCACCAACCAATGACGCTGTTTACAGAGCGGCATGTACGAACCGCGAGAGCTTGT 1195
Db 541 GTTCCACCAACCAATGACGCTGTTTACAGAGCGGCATGTACGAACCGCGAGAGCTTGT 600
QY 1196 TGAGCCAGATTTTCATATAGTGAAGAGTGTGTTGTCCTTCATATTTGCAAG 1255
Db 601 TGAGCCAGATTTTCATATAGTGAAGAGTGTGTTGTCCTTCATATTTGCAAG 660
QY 1256 ACCACAAATGAGCTAAGATTGTACTGTTTTCAGATTCTAT--CGATTTCATTATGAAAA 1314
Db 661 ACCACAAATGAGCTAAGATTGTACTGTTTTCAGATTCTATTTTATGAAAA 720

```

QY 1315 CTGTGTTGCCACAGTAGAATGTC-TGTGAACAGAGAGACCCCTGTGTGGTTCATGCTAAC 1373
Db 721 CTGTGTTGCCACAGTAGAATGTC-TGTGAACAGAGAGAGACCCCTGTGTGGTTCATGCTAAC 780
QY 1374 AAGACACAAAGTCTGTCTTCCCTGAACC-ATGTGGATACTTTACAGA--AATGGACTGG 1430
Db 781 AAGACACAAAGTCTGTCTTCCCTGAACCAATGTGAATTAATTTTACAGACATGGACCTGG 840
QY 1431 AGTCATCTGCAAAAGCCCTCTTGTAAAGACTGGTTTC-TGCCAATGACCAACAGCCCA 1489
Db 841 AGTCATCTGCAAAAGCCCTCTTGTAAAGACCCTTTTCTTGCCATTGACCAACAGCCCC 900
QY 1490 AGATTTTCCCTCTGTGATTTCTTTAAAGAAATGACTATATATATTTTCCACTTAAAT 1549
Db 901 AGATTTTCCCTCTGTGATTTCTTTAAAGAAATGACTATATATATTTTCCACTTAAAT 1549
QY 1550 ATTGTTTCTGCATTCATTTTATAGCAACCAACAATTG 1586
Db 958 TTGTTCTGCAATCAATTTTATAGCAACCAACAATTG 994

RESULT 3
BG680919 831 bp mRNA linear EST 01-MAY-2001
LOCUS 602628716F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4753755 5',
DEFINITION mRNA sequence.
ACCESSION BG680919
VERSION BG680919.1 GI:13912303
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 831)
AUTHORS NIH-MGC http://www.ncbi.nlm.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM10613 row: o column: 04
High quality sequence stop: 810.
Location/Qualifiers
1. 831
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4753755"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Skn4"
/notes="Organ: skin; Vector: pCMV-Sport6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

ORIGIN
Query Match 46.7%; Score 782.4; DB 12; Length 831;
Best Local Similarity 99.4%; Pred. No. 2.5e-118;
Matches 827; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

QY 43 GTTCTCTGTCGGCGTGCCTGCTCCGGGTCTCGAGAGCGCCCGCGCGCGCGCG 102
Db 1 GTTCTCTGTCGGCGTGCCTGCTCCGGGTCTCGAGAGCGCCCGCGCGCGCGCG 60
QY 103 CTTTCGAGTCCGACTCGACCTCTCGAGCGCGAGCCCGAGCGCGCGAGCCAGGCTT 162

```

```

Db 61 CTTTCGAGTCCGGAATCGACCTCTCGAGCGCGAGCCCGAGCGCGAGCCAGCGTT 120
QY 163 ATGCAAGCAAAAGATCTGGAGGAGCAGTTTACGGTCTGTGTCCAGTGTAGTCAATCATGA 222
Db 121 ATGCAAGCAAAAGATCTGGAGGAGCAGTTTACGGTCTGTGTCCAGTGTAGTCAATCATGA 180
QY 223 CTGTACTCTACCCAGATATTGGAAATGTACAAGTGTCAAGTGTCAAGTAAGGAAGAGGCTGGC 282
Db 181 CTGTACTCTACCCAGATATTGGAAATGTACAAGTGTCAAGTGTCAAGTAAGGAAGAGGCTGGC 240
QY 283 AACATACAGAGAACAGCCCAACCTCAACTCAAGCAGAGAGAGACTATATAAATTTGCTG 342
Db 241 AACATACAGAGAACAGCCCAACCTCAACTCAAGCAGAGAGAGACTATATAAATTTGCTG 300
QY 343 CAGCACATTATAATACAGAGATCTTGAAGATTTGATAATGAGTGGAGAAAGACTCAAT 402
Db 301 CAGCACATTATAATACAGAGATCTTGAAGATTTGATAATGAGTGGAGAAAGACTCAAT 360
QY 403 GCATGCCACGGAGGTGTGTATAGATGTGGGAGGAGTTTGGAGTCCGCAACACACCT 462
Db 361 GCATGCCACGGAGGTGTGTATAGATGTGGGAGGAG- TTGGAGTCCGCAACACACCT 419
QY 463 TCTTTAAACCTCCATGTGTGTCCGTCTACAGATGTGGGGTTGTGCAATAGTAGGGGC 522
Db 420 TCTTTAAACCTCCATGTGTGTCCGTCTACAGATGTGGGGTTGTGCAATAGTAGGGGC 479
QY 523 TGCAGTGCATGAACACAGACAGAGTACCTCAGCAAGAGCTTATTTTGAATTTACAGTGC 582
Db 480 TGCAGTGCATGAACACAGACAGAGTACCTCAGCAAGAGCTTATTTTGAATTTACAGTGC 539
QY 583 CTCTCTCTCAAGGCCCCCAACACCTCACTCACTTTTGGCAATCACACTTCCCTGCCGAT 642
Db 540 CTCTCTCTCAAGGCCCCCAACACCTCACTCACTTTTGGCAATCACACTTCCCTGCCGAT 599
QY 643 GCATGTCTAAACTGGATGTTTACAGACAGTTCATTCATTTATAGAGCTTCCCTGCCAG 659
Db 600 GCATGTCTAAACTGGATGTTTACAGACAGTTCATTCATTTATAGAGCTTCCCTGCCAG 628
QY 703 CAACACTACACAGTGTCCAGCAGCGAACAGACCTCCGCCCAACCAATATAGTGGGATA 762
Db 660 CAACA-TACACAGTGTCCAGCAGCGAACAGACCTCCGCCCAACCAATATAGTGGGATA 718
QY 763 ATCATCTCTCAGATGCTGCTCAGGAAGATTTATGTTTCTCGGATGCTCGAGATG 822
Db 719 ATCATCTCTCAGATGCTGCTCAGGAAGATTTATGTTTCTCGGATGCTCGAGATG 777
QY 823 ACTCAACAGATGGATTCCATGACATCTGTGACCAACAAAGAGAGCTGGATGA 874
Db 778 ACTCAACAGATGGATT-CATGACATCTGTGACCAACAAAGAGAGCTGGATGA 828

```

RESULT 4

CA488579

LOCUS

DEFINITION

AGENCY

AGENCY

AGENCY

AGENCY

AGENCY

AGENCY

AGENCY

AGENCY

AGENCY

AGENCY

AGENCY

AGENCY

AGENCY

AGENCY

AGENCY

AGENCY

AGENCY

AGENCY

AGENCY

AGENCY

AGENCY

AGENCY

AGENCY


```
Query Match      43.6%; Score 730.6; DB 14; Length 750;
Best Local Similarity 99.2%; Pred. No. 7.9e-110;
Matches 744; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 925 GACCCCAAGAACTAGACAGAACTCATGCGAGTGTGCTGTATATAAACAACACTCTTCC 984
DB 750 GACCCCAAGAACTAGACAGAACTCATGCGAGTGTGCTGTATATAAACAACACTCTTCC 992
QY 985 CAGCCCAATGTGGGCGCAACCGAGAAATTTGATGAAACACATCCAGTGTGTATGATAA 1044
DB 691 CAGCCCAATGTGGGCGCAACCGAGAAATNTGATGAAACACATCCAGTGTGTATGATAA 632
QY 1045 GAACCTGCCCGAGAAATCAACCCCTTAATCTCGAAAGTGTGCTGTGAATGTACAGAA 1104
DB 631 GAACCTGCCCGAGAAATCAACCCCTTAATCTCGAAAGTGTGCTGTGAATGTACAGAA 572
QY 1105 GTCCACAGAAATGCTTTTAAAGAGAAAGTTCACACCAACCAACATGAGCTGTATCA 1164
DB 571 GTCCACAGAAATGCTTTTAAAGAGAAAGTTCACACCAACCAACATGAGCTGTATCA 512
QY 1165 GAGGCCATATGACGACCGCAGAGCTTGTGAGCGCAGGATTTTCATATAGTGAAGAAG 1224
DB 511 GAGGCCATATGACGACCGCAGAGCTTGTGAGCGCAGGATTTTCATATAGTGAAGAAG 452
QY 1225 TGTGCTGCTGTGCTCCCTTCATATTGCAAGACCAACCAATGAGCTAAGATTGTACTGTTT 1284
DB 451 TGTGCTGCTGTGCTCCCTTCATATTGCAAGACCAACCAATGAGCTAAGATTGTACTGTTT 392
QY 1285 TCCAGTTCATCGAATTTCTATTATGAAACACTGTGTCACAGTGAAGACTGTCTGTGAA 1344
DB 391 TCCAGTTCATCGAATTTCTATTATGAAACACTGTGTCACAGTGAAGACTGTCTGTGAA 332
QY 1345 CAGAGAGACCTTGTGGTGCATGCTAAACAAACAGAAAGTGTCTTTTCCGTAACCATG 1404
DB 331 CAGAGAGACCTTGTGGTGCATGCTAAACAAACAGAAAGTGTCTTTTCCGTAACCATG 272
QY 1405 TGGATAACTTTACAGAAATGAGTGGAGTCACTGCAAAAGGCCCTTTGTAAAGACTGG 1464
DB 271 TGGATAACTTTACAGAAATGAGTGGAGTCACTGCAAAAGGCCCTTTGTAAAGACTGG 212
QY 1465 TTTTCTGCCAATGACCAACAGCCAGATTTTCTCTCTGTGATTTTAAAGATGAC 1524
DB 211 TTTTCTGCCAATGACCAACAGCCAGATTTTCTCTCTGTGATTTTAAAGATGAC 152
QY 1525 TATATAATTTATTTCCATAAAATATGTTTCTGCAATTCATTTTATAGCAACCAAT 1584
DB 151 TATATAATTTATTTCCATAAAATATGTTTCTGCAATTCATTTTATAGCAACCAAT 92
QY 1585 TGTATAAACTCAGTGTGATCAATATTTTATATCATGCAAAATATGTTTAAATAAATG 1644
DB 91 TGTATAAACTCAGTGTGATCAATATTTTATATCATGCAAAATATGTTTAAATAAATG 32
QY 1645 AAAATTGTATTTATAAAAAAATAAAAAA 1674
DB 31 AAAATTGTATTTATAAAAAAATAAAAAA 2

RESULT 6
BU615444/c
LOCUS      748 bp mRNA linear EST 23-SEP-2002
DEFINITION UI-H-FHO-bco-p-03-0-UI.s1 NCI CGAP FHO Homo sapiens cDNA clone
ACCESSION  BU615444
VERSION     BU615444.1 GI:23281659
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 748)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
```

JOURNAL
COMMENT

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: James Martin
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@iowa.edu
The following repetitive elements were found in this cDNA
sequence: 1-51, >AT rich#Low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA-Yes.

FEATURES

Location/Qualifiers
source

1..748
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FHO-bco-p-03-0-UI"
/tissue_type="Human Chondrosarcoma Cell Line"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP FHO"
/note="Organ: Bone; Vector: p77T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI CGAP FHO is a cDNA library containing the following
tissue(s): Human Grade 1 Chondrosarcoma Cell line The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into p77T3-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dT)18 tail. The sequence tag for this
library is AGAATCCGCG. The cell line was provided by Dr
James Martin from University of Iowa
TAG TISSUE=Human Chondrosarcoma Cell Line CS8 - Grade 1
Chondrosarcoma
TAG LIB=UI-H-FHO
TAG_SEQ=AGAATCCGCG"

ORIGIN

Query Match 43.6%; Score 729.6; DB 13; Length 748;
Best Local Similarity 99.2%; Pred. No. 1.1e-109;
Matches 743; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
QY 925 GACCCCAAGAACTAGACAGAACTCATGCGAGTGTGCTGTATATAAACAACACTCTTCC 984
DB 748 GACCCCAAGAACTAGACAGAACTCATGCGAGTGTGCTGTATATAAACAACACTCTTCC 689
QY 985 CCAGCCCAATGTGGGCGCAACCGAGAAATTTGATGAAACACATGCGAGTGTGTATGATAA 1044
DB 588 CCAGCCCAATGT-GGGCCCAACCGAGAAATNTGATGAAACACATGCGAGTGTGTATGATAA 630
QY 1045 GAACCTGCCCGAGAAATCAACCCCTTAATCTCGAAAGTGTGCTGTGAATGTACAGAA 1104
DB 629 GAACCTGCCCGAGAAATCAACCCCTTAATCTCGAAAGTGTGCTGTGAATGTACAGAA 570
QY 1105 GTCCACAGAAATGCTTTTAAAGAGAAAGTTCACACCAACCAACATGAGCTGTATCA 1164
DB 569 GTCCACAGAAATGCTTTTAAAGAGAAAGTTCACACCAACCAACATGAGCTGTATCA 510
QY 1165 GAGGCCATGTACGAACCGCAGAGCTTGTGAGCGCAGGATTTTCATATAGTGAAGAAG 1224
DB 509 GAGGCCATGTACGAACCGCAGAGCTTGTGAGCGCAGGATTTTCATATAGTGAAGAAG 450
QY 1225 TGTGCTGCTGTGCTCCCTTCATATTGCAAGACCAACCAATGAGCTAAGATTGTACTGTTT 1284
DB 449 TGTGCTGCTGTGCTCCCTTCATATTGCAAGACCAACCAATGAGCTAAGATTGTACTGTTT 390
QY 1285 TCCAGTTCATCGAATTTCTATTATGAAACACTGTGTCACAGTGAAGACTGTCTGTGAA 1344

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 (bases 1 to 717)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 785 Std Error: 0.00
 Seq primer: -40ml3 fwd. EF from Amersham
 High quality sequence stop: 476.
 Location/Qualifiers
 1. 717
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1544257"
 /lab_host="DH10B"
 /clone_lib="Soares_NFL_T_GBC_S1"
 /notes="Organ: pooled; Vector: pVT73D-Pac (Pharmacia) with
 a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
 Equal amounts of plasmid DNA from three normalized
 libraries (fetal lung NBHL19W, testis NHT, and B-cell
 NCI CGAP GCBI) were mixed, and ss circles were made in
 vitro. Following HAP purification, this DNA was used as
 tracer in a subtractive hybridization reaction. The driver
 was PCR-amplified cDNAs from pools of 5,000 clones made
 from the same 3 libraries. The pools consisted of
 I.M.A.G.E. clones 297480-302087, 682632-687239,
 726408-728711, and 729096-731399. Subtraction by Bento
 Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 41.7%; Score 698.6; DB 9; Length 717;
 Best Local Similarity 99.6%; Pred. No. 1.4e-104;
 Matches 704; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 948 AACCTCATGCTGTGCTGTAAACAACTCTTCCACGCAATGTGGGGCAACCGA 1007
 Db 714 AATCATGCTGTGCTGTAAACAACTCTTCCACGCAATGTGGGGCAACCGA 655

QY 1008 GAATTTGATGAACACATCCAGTGTGTATGTAAGAAGACCTGCCCGAATCAACCC 1067
 Db 654 GAATTTGATGAACACATCCAGTGTGTATGTAAGAAGACCTGCCCGAATCAACCC 595

QY 1068 CTAAATCTGGAAATGTGCTGTGAATGTACAGAAATCCAGAAATGCTTTTAAAA 1127
 Db 594 CTAAATCTGGAAATGTGCTGTGAATGTACAGAAATCCAGAAATGCTTTTAAAA 535

QY 1128 GGAAGAAGATTCCACCACCAACATGCTGTGTACAGAGCGCATGTACGACCCCGAG 1187
 Db 534 GGAAGAAGATTCCACCACCAACATGCTGTGTACAGAGCGCATGTACGACCCCGAG 475

QY 1188 AAGCTTTGTAGCCAGAGATTTCATATAGTGAAGAGTGTGTGTGTGTCCTTCATAT 1247
 Db 474 AATGCTTTGTAGCCAGAGATTTCATATAGTGAAGAGTGTGTGTGTGTCCTTCATAT 415

QY 1248 TGGCAAGACCCACAAATGAGTAAAGTGTACTGTTTCCAGTTTCAGTTTTCATTAT 1307
 Db 414 TGGCAAGACCCACAAATGAGTAAAGTGTACTGTTTCCAGTTTCAGTTTTCATTAT 355

QY 1308 TGGAAATCTGTGTGCCACAGTAGAATGTGTGTGAACAGAGAGACCTTGTGGTCCAT 1367
 Db 354 TGGAAATCTGTGTGCCACAGTAGAATGTGTGTGAACAGAGAGACCTTGTGGTCCAT 295

QY 1368 GCTAACAAAGACAAAGTCTGCTTCTTCCATGACCATGTGTGATACCTTTACAGAAATGGAC 1427
 Db 294 GCTAACAAAGACAAAGTCTGCTTCTTCCATGACCATGTGTGATACCTTTACAGAAATGGAC 235

QY 1428 TGGAGCTCATGTGCAAAAGGCGCTCTTGTAAAGACTGTGTTTCTGCCAATGACCAACAGC 1487

Db 234 TGGAGCTCATGTGCAAAAGGCGCTCTTGTAAAGACTGTGTTTCCCAATGACCAACAGC 175
 QY 1488 CAAGATTTTCTCTTGTGATTCTTTAAAGAAAGTACTATATATATTTATTTTCCACTAAAA 1547
 Db 174 CAAGATTTTCTCTTGTGATTCTTTAAAGAAAGTACTATATATATTTATTTTCCACTAAAA 115
 QY 1548 ATATGTTTCTGCATTCATATTTTATAGCAACAAATGGTAAACCTCCTGTGATCAAT 1607
 Db 114 ATATGTTTCTGCATTCATATTTTATAGCAACAAATGGTAAACCTCCTGTGATCAAT 55
 QY 1608 ATTTTATATCATGCAAAATATGTTTAAATAAATAAATGAATAATCTATTATAAA 1661
 Db 54 ATTTTATATCATGCAAAATATGTTTAAATAAATAAATGAATAATCTATTATAAA 1

RESULT 9
 AWO70215/c
 LOCUS AWO70215 728 bp mRNA linear EST 20-OCT-2000
 DEFINITION xa09c10.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
 IMAGE:2567826.3, similar to SW:VEGC_HUMAN P49767 VASCULAR
 ENDOTHELIAL GROWTH FACTOR C PRECURSOR 1, mRNA sequence.
 ACCESSION AWO70215 Gi:6025213
 VERSION AWO70215
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 728)
 REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 2601 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 457.
 Location/Qualifiers
 1. 728
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2567826"
 /lab_host="DH10B"
 /clone_lib="Soares_NFL_T_GBC_S1"
 /notes="Organ: pooled; Vector: pVT73D-Pac (Pharmacia) with
 a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
 Equal amounts of plasmid DNA from three normalized
 libraries (fetal lung NBHL19W, testis NHT, and B-cell
 NCI CGAP GCBI) were mixed, and ss circles were made in
 vitro. Following HAP purification, this DNA was used as
 tracer in a subtractive hybridization reaction. The driver
 was PCR-amplified cDNAs from pools of 5,000 clones made
 from the same 3 libraries. The pools consisted of
 I.M.A.G.E. clones 297480-302087, 682632-687239,
 726408-728711, and 729096-731399. Subtraction by Bento
 Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 39.9%; Score 667.2; DB 9; Length 728;
 Best Local Similarity 97.7%; Pred. No. 1.9e-99;
 Matches 708; Conservative 0; Mismatches 14; Indels 3; Gaps 3;

QY 933 AAAGAACTAGACAGAACTCATGCCAGTGTCTGTAAACAACTCTTCCCGAGCAA 992
 Db 728 AAAGAACTAGACAGAACTCATGCCAGTGTCTGTAAACAACTCTTCCCGAGCAA 670
 QY 993 TGTGGGGCCAAACCGAGATTTGATGAAACACATGCCAG-TGTGTATGTAAGAACTGT 1051

sequence: 1-53, >AT rich#Low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
source

Location/Qualifiers
1..662
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="IMAGE:5842762"
/tissue_type="Chondrosarcoma"
/dev_stage="Adult"
/lab_hosts="DH10B (Life Technologies)"
/clone_lib="NCI CGAP ED0"
/note="Organ: Left Public Bone; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP ED0 is a cDNA library containing the following tissue(s): Chondrosarcoma cell line C55. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GCTCAAGGCT.
TAG_TISSUE=chondrosarcoma
TAG_LIB=UI-H-ED0
TAG_SEQ=CGTCAAGGCT"

ORIGIN

Query Match 39.2%; Score 556.2; DB 12; Length 662;
Best Local Similarity 99.5%; Pred. No. 1.2e-97;
Matches 658; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1014 GATCAAAACACATGCCAGTGTGTATGTAAGAAGACCTGCCCCAGAAATCAACCCCTAAAT 1073
DB 662 GATGAAACACATGCCAGTGTGTATGTAAGAAGACCTGCCCCAGAAATCAACCCCTAAAT 603
QY 1074 CTTGAAAAATGTGCTGTGAATGTACAGAAAGTCCACAGAAATCTGTGTTAAAGAGAAAG 1133
DB 602 CTTGAAAAATGTGCTGTGAATGTACAGAAAGTCCACAGAAATCTGTGTTAAAGAGAAAG 543
QY 1134 AAGTTCCACCAACCAACATGCAGCTGTACAGAGCGCATGTACGAAACCGCAGAGGCT 1193
DB 542 AAGTTCCACCAACCAACATGCAGCTGTACAGAGCGCATGTACGAAACCGCAGAGGCT 483
QY 1194 TGTGAGCCGAGATTTTCATATAGTGAAGAAGTGTGCTGTGTGTCCTTCATATTTGCAA 1253
DB 482 TGTGAGCCGAGATTTTCATATAGTGAAGAAGTGTGCTGTGTGTCCTTCATATTTGCAA 423
QY 1254 AGACCAACAATGAGCTAAGATTGTACTGTTTCCAGTTTCATCGATTTTCTATTTATGAAA 1313
DB 422 AGACCAACAATGAGCTAAGATTGTACTGTTTCCAGTTTCATCGATTTTCTATTTATGAAA 363
QY 1314 ACTGTGTTGCCACAGTAGAACTGTCTGTAACACAGAGAGCCCTGTGGTCCATGCTAAC 1373
DB 362 ACTGTGTTGCCACAGTAGAACTGTCTGTAACACAGAGAGCCCTGTGGTCCATGCTAAC 303
QY 1374 AAAGCAAAAAGTGTGCTTTCTCTGAAACATGTGTGATTAACCTTTACAGAAATGAGCTGAGC 1433
DB 302 AAAGCAAAAAGTGTGCTTTCTCTGAAACATGTGTGATTAACCTTTACAGAAATGAGCTGAGC 243
QY 1434 TCATCTGCAAAAGGCTCTTGTGAAGAGTGTGTTTCTGCCAATGACCAACAGCAAGAT 1493
DB 242 TCATCTGCAAAAGGCTCTTGTGAAGAGTGTGTTTCTGCCAATGACCAACAGCAAGAT 183
QY 1494 TTTCTCTGTGATTTCTTTAAAGAGTACTATATATTTATTTTCCATTAATAATTTG 1553
DB 182 TTTCTCTGTGATTTCTTTAAAGAGTACTATATATTTATTTTCCATTAATAATTTG 123
QY 1554 TTTCTGCTTCATTTTATAGCAACAATTTGTTAAACCTCACTGTGATCAATATTTT 1613

669 GTGGGGCAACCCGAGATTGNTGATGAACACATGCGCAGNTGTTGATGTAAGAAGACCTG 610
1052 CCCAGAAATCAACCCCTAAATCTCGAAATGTGCTGTGAATGTACAGAAAGTCCACA 1111
609 CCCAGAAATCAACCCCTAAATCTCGAAATGTGCTGTGAATGTACAGAAAGTCCACA 550
1112 GAAATGCTTTTAAAGGAAGAGTTCACACCAACCAACATGAGCTGTACAGAGGCG 1171
549 GAAATGCTTTTAAAGGAAGAGTTCACACCAACCAACATGAGCTGTACAGAGGCG 490
1172 ATGTACGAACCGCGCAGAGGCTGTGAGCCAGGATTTTCATATAGTGAAGAAGTGTGCG 1231
489 ATGTACGAACCGCGCAGAGGCTGTGAGCCAGGATTTTCATATAGTGAAGAAGTGTGCG 430
1232 TTGTGTCCTTCATATTTGGCAAGACCAACCAACCAACATGAGCTGTGTTTCCAGTT 1291
429 TTGTGTCCTTCATATTTGGAAAGACCAACCAACCAACATGAGCTGTGTTTCCAGTT 370
1292 CATCGATTTTCTATTTATGGAAGAACTGTGTTGCCACAGTAGAACTGTGTAACAGAG 1351
369 CATCGATTTTCTATTTATGGAAGAACTGTGTTGCCACAGTAGAACTGTGTAACAGAG 310
1352 ACCCTTGTGGGTCATCTAACAAAGACCAAAAGTCTGCTTTCTGAAACCATGTGGATAA 1411
309 ACCCTTGTGGTCCATCTAACAAAGACCAAAAGTCTGCTTTCTGAAACCATGTGGATAA 250
1412 CTTTACGAATGAGCTGG-AGCTCATCTGCAAAAGGCTCTTGTGAAGACGTGTTTCT 1470
249 CTTTACGAATGAGCTGGAAGCTCATCTGCAAAAGGCTCTTGTGAAGACGTGTTTCT 190
1471 GCCAATGACCAACAGCAAGATTTTCTCTGTGATTTCTTTAAAGAAATGACTATATA 1530
189 GCCAATGACCAACAGCAAGATTTTCTCTGTGATTTCTTTAAAGAAATGACTATATA 130
1531 ATTATTTCCACTAAAATATTGTTCTGCAATTCATTTTATAGCAACAACCAATTTGTTAA 1590
129 ATTATTTCCACTAAAATATTGTTCTGCAATTCATTTTATAGCAACAACCAATTTGTTAA 70
1591 AACTCACTGTGATCAATATTTTATATCATGCAAAATATGTTTAAATTAATGAAATTT 1650
69 AACTCACTGTGATCAATATTTTATATCATGCAAAATATGTTTAAATTAATGAAATTT 10
1651 GTATT 1655
9 GTATT 5

RESULT 10
BQ005831/c
LOCUS
DEFINITION
IMAGE:5842762 3', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
EST.
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 662)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-x@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
The following repetitive elements were found in this cDNA

```

Db      122 TTTCGCAATCAATTTATAGCAACAACAATGGTAAACCTCACTGTGATCAATATTTT 63
1614 ATATCATGCAAAATATGTTTAAATAAATAAATGAAATTCGATTATTAATAAATAAATAAATAA 1673
Db      62 ATATCATGCAAAATATGTTTAAATAAATAAATGAAATTCGATTATTAATAAATAAATAA 3
Qy      1674 A 1674
Db      2 A 2

RESULT 11
LOCUS   BU619991/c
DEFINITION   UI-H-FL1-bfv-p-05-0-UI.s1 NCI CGAP FL1 Homo sapiens cDNA clone
ACCESSION   BU619991
VERSION     BU619991.1 GI:23286206
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 657)
AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL    Unpublished (1997)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Martin
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
            The following repetitive elements were found in this cDNA
            sequence: 1-48, >AT rich#Low complexity (matched complement)
            Seq primer: M13 FORWARD
            POLYA=Yes

FEATURES             location/Qualifiers
     1..657
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="UI-H-FL1-bfv-p-05-0-UI"
     /tissue_type="Cell lines"
     /dev_stage="Adult"
     /lab_host="DH10B (Life Technologies)"
     /clone_lib="NCI CGAP FL1"
     /note="Organ: Chondrosarcoma; Vector: pTTT3-Pac
            (Pharmacia) with a modified polylinker; Site 1: EcoR I;
            Site 2: Not I; NCI CGAP FL1 is a normalized cDNA library
            derived from a pool of mRNA obtained from 4 cell lines
            from grade III chondrosarcoma tissues. The library was
            constructed according to Bonaldo, Lennon and Soares,
            Genome Research, 6:791-806, 1996. First strand cDNA
            synthesis was primed with an oligo-dT primer containing a
            Not I site. Double stranded cDNA was ligated to an EcoR I
            adaptor, digested with Not I, and cloned directionally
            into pTTT3-Pac vector. The oligonucleotide used to prime
            the synthesis of first-strand cDNA contains a library tag
            sequence that is located between the Not I site and the
            (dT)18 tail. The sequence tag for this library is
            GAGGTCGGTG. The cell lines were provided by Dr. James
            Martin from the University of Iowa.
            TAG TISSUE=Human Chondrosarcoma Grade 3 cell line mix
            TAG_LIB=UI-H-FL1
            TAG_SEQ=GAGGTCGGTG"

ORIGIN
Query Match      38.8%; Score 649; DB 13; Length 657;

Best Local Similarity 99.2%; Pred. No. 1.8e-96;
Matches 652; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

1014 GATGAAAACACATGCCAGTGTGTATGTAAGAAACCTGCCAGAAATCAACCCCTAAAT 1073
Db      657 GATGAAAACACATGCCAGTGTGTATGTAAGAAACCTGCCAGAAATCAACCCCTAAAT 598
Qy      1074 CTTGSAAAATGTGCTGTGAATGTACAGAAAGTCCAAGAAATGCTTGTAAAGGAAAG 1133
Db      597 CTTGSAAAATGTGCTGTGAATGTACAGAAAGTCCAAGAAATGCTTGTAAAGGAAAG 538
Qy      1134 AAGTTCCACCAACCAATGCAGCTGTTACAGACGGCCATGTACGACCGCAGAGGCT 1193
Db      537 AAGTTCCACCAACCAATGCAGCTGTTACAGACGGCCATGTACGACCGCAGAGGCT 478
Qy      1194 TGTAGCCAGGAGTATTTTCATATAGTGAAGAGTGTGTGTTGTGTCCTTCATATGGCAA 1253
Db      477 TGTAGCCAGGAGTATTTTCATATAGTGAAGAGTGTGTGTTGTGTCCTTCATATGGCAA 418
Qy      1254 AGACCAAAATGAGCTAAGATTGTACTGTTTCCAGTTCATCGATTTCTATATGGAAA 1313
Db      417 AGACCAAAATGAGCTAAGATTGTACTGTTTCCAGTTCATCGATTTCTATATGGAAA 358
Qy      1314 ACTGTGTGTCACAGTAGAACTGTCTGCAACAGAGAGACCCCTGTGGGTCCATGCTAAC 1373
Db      357 ACTGTGTGTCACAGTAGAACTGTCTGCAACAGAGAGACCCCTGTGGGTCCATGCTAAC 298
Qy      1374 AAAGCAAAAGTGTCTTTCTCTGAAACCAATGGGATAACTTTACAGAAATGGAAGTGGAGC 1433
Db      297 AAAGCAAAAGTGTCTTTCTCTGAAACCAATGGGATAACTTTACAGAAATGGAAGTGGAGC 238
Qy      1434 TCATCTGCAAAAGGCTCTTTGTAAGACCTGTTTCTGCAATGACCAACAGCCCAAGAT 1493
Db      237 TCATCTGCAAAAGGCTCTTTGTAAGACCTGTTTCTGCAATGACCAACAGCCCAAGAT 178
Qy      1494 TTTCTCTCTTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAAAATATTG 1553
Db      177 TTTCTCTCTTGTGATTTCTTTAAAGAAATGACTATATAATTTATTTCCACTAAAAATATTG 118
Qy      1554 TTTCTGCAATTTTATAGCAACCAATTTGGTAAACTCACTGTGATCAATATTTT 1613
Db      117 TTTCTGCAATTTTATAGCAACCAATTTGGTAAACTCACTGTGATCAATATTTT 58
Qy      1614 ATATCATGCAAAATATGTTTAAATAAATAAATGTAATTTATTAATAAATAAATAAATAA 1670
Db      57 ATATCATGCAAAATATGTTTAAATAAATAAATGTAATTTATTAATAAATAAATAAATAA 1

RESULT 12
LOCUS   BE218812/c
DEFINITION   BE218812 637 bp mRNA linear EST 03-JUL-2000
            hv45a10.x1 NCI CGAP Lu24 Homo sapiens cDNA clone IMAGE:3176346 3,
            similar to SW:VEGC HUMAN P49767 VASCULAR ENDOTHELIAL GROWTH FACTOR
            C PRECURSOR ;, mRNA sequence.
ACCESSION   BE218812
VERSION     BE218812.1 GI:8906130
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 637)
AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL    Unpublished (1997)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center

```

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 449.

```

FEATURES
source
  Location/Qualifiers
    1..637
      /organism="Homo sapiens"
      /mcl_type="RNA"
      /db_xref="taxon:9606"
      /clone_image="3176346"
      /issue_type="carcinoid"
      /lab_host="DH10B"
      /clone_lib="NCI_CGAP_Lu2"
      /note="Organ: Lung; Vect
modified polylinker; Pla
library NCI_CGAP_Lu2 was
made in vitro. Following
used as tracer in a sub
The driver was PCR-ampl
clones made from the sam
1414920-1417991 and 1520
Soares and M. Patricia

```

Query Match	36.8%	Score 616	DB 10	Length 637
Best Local Similarity	99.1%	Pred. 4.7e-91		
Matches 630	Conservative 0	Mismatches 5	Indels 1	Gaps 1
QY	1020	AAACATGCCAGTGTGATGTAAAGAGACCTGCCCCAGAAATCAACCCCTAAATCTCTGGA	1079	
DB				
QY	637	AAACATGCCAGTGTGTATGT-AAAGAACTTGCCCGAGAAATCACCCCTTAAATCTCTGGA	579	
DB				
QY	1080	AAATGTGCTCTGGAATGTACAGAAAGTCCACAGAAATGCTTGTTAAAGGAAAGAAAGTTC	1139	
DB				
QY	578	AAATGTGCTTGTGAATGTACAGAAAGTCCACAGAAATGCTTGTTAAAGGAAAGAAAGTTC	519	
DB				
QY	1140	CACCACCAACATGCAGCTGTTACAGACGCCCATGTACGAACCCCGAGAGGCTTGTGAG	1199	
DB				
QY	518	CACCACCAACATGCAGCTGTTACAGACGCCCATGTACGAACCCCGAGAGGCTTGTGAG	459	
DB				
QY	1200	CCAGGATTTTCATATAGTAGAAGAGTGTGCTGTGCTCCCTTCATATTTGGCAAGACCA	1259	
DB				
QY	458	CCAGGATTTTCATATAGTAGAAGAGTGTGCTGTGCTCCCTTCATATTTGGAAAGACCA	399	
DB				
QY	1260	CAAAATGAGCTAAAGATGTACTGTTTTCCAGTTTCATCGATTTTCTATTATGGAAGAACTGTG	1319	
DB				
QY	398	CAAAATGAGCTAAAGATGTACTGTTTTCCAGTTTCATCGATTTTCTATTATGGAAGAACTGTG	339	
DB				
QY	1320	TTGCCACAGTAGAACTGCTGTGTAACAGAGAGACCCCTTGTGGGTCCATGCTTAAACAAAGAC	1379	
DB				
QY	338	TTGCCACAGTAGAACTGCTGTGTAACAGAGAGACCCCTTGTGGGTCCATGCTTAAACAAAGAC	279	
DB				
QY	1380	AAAAGTCTGTCTTTCCTGAACCATGTGGATAACCTTACAGAAATGGAGCTGCATCT	1439	
DB				
QY	278	AAAAGTCTGTCTTTCCTGAACCATGTGGATAACCTTACAGAAATGGAGCTGCATCT	219	
DB				
QY	1440	GCAAAAGGCTCTTGTTAAAGACTGGTTTTCTGCCAATGACCACACAGCAAGATTTTCCT	1499	
DB				
QY	218	GCAAAAGGCTCTTGTTAAAGACTGGTTTTCTGCCAATGACCACACAGCAAGATTTTCCT	159	
DB				
QY	1500	CTTGTGATTTCTTTTAAAGAAAGTACTATATAATTTATTTCCACTAAAAATATTTGTTCTG	1559	
DB				
QY	158	CTTGTGATTTCTTTTAAAGAAAGTACTATATAATTTATTTCCACTAAAAATATTTGTTCTG	99	
DB				
QY	1560	CATTCAATTTTATAGCAACCAATTTGGTAAAACTCACTGTGATCAATATTTTTATATCA	1619	
DB				
QY	98	CATTCAATTTTATAGCAACCAATTTGGTAAAACTCACTGTGATCAATTTTTTATATCA	39	
DB				
QY	1620	TGCAAAATATGTTTTAAAAATAAAATGAAATTTGTATT	1655	
DB				
QY	38	TGCAAAATATGTTTTAAAAATAAAATGAAATTTGTATT	3	
DB				

```

RESULT 13
CB854820/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

CB854820
UI-CF-EC1-ack-o-03-0-UI.s1
UI-CF-EC1-ack-o-03-0-UI.3', mRNA sequence.
CB854820
CB854820.1
GI:30045199
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 622)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalisation and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
The following repetitive elements were found in this cDNA
sequence: 1-53, >AtrichFlow_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
Location/Qualifiers
1..622
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-EC1-ack-o-03-0-UI"
/tissue_type="Lung"
/dev_stage="Adult and Fetal"
/lab_host="DH10B (Life Technologies)" (T1 phage resistant)"
/clone_lib="UI-CF-EC1"
/notes="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-CF-EC1 is a normalized cDNA library containing the
following tissue(s): Normal lung from adult and from fetal
day 64, day 87, week 19 and week 42. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT73-pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
AAGTGGCTTAC.
TAG SEO=None found"

```

ORIGIN		100% Bootstrap	
Query Match	36.7%	Score 614.6;	DB 14; Length 622;
Best Local Similarity	99.4%	Pred. No. 7.9e-91;	
Matches 617;	Conservative 0;	Mismatches 4;	Indels 0; Gaps 0;
Qy	1054	CCAGAAATCAACCCCTAAATCTCTGGAAAATGTCCCTGTGAATGTACAGAAAGTCCACAGA	1113
Db	622	CCAGAAATCAACCCCTAAATCTCTGGAAAATGTCCCTGTGAATGTACAGAAAGTCCACAGA	563

RESULT 15
 BM1271186/c
 LOCUS
 DEFINITION
 BM1271186 584 bp mRNA linear EST 12-MAR-2002
 ie95f08.x1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
 cDNA clone IMAGE:5674886 3' similar to SW:VEGC_HUMAN P49767
 VASCULAR ENDOTHELIAL GROWTH FACTOR C PRECURSOR , mRNA sequence.
 BM1271186
 BM1271186.1 GI:17121738
 EST.
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 SOURCE
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (Bases 1 to 584)
 AUTHORS
 Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
 Lemishka,I., Szeacrae,W., Brestelli,J., Gradwohl,G., Clifton,S.,
 Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blustein,A.,
 Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
 Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
 Williams,T., Jackson,Y. and Bowers,Y.
 TITLE
 Endocrine Pancreas Consortium
 JOURNAL
 Unpublished (2000)
 COMMENT
 Other_ESRs: ie95f08.y1
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@biohph.harvard.edu
 Library was constructed by Dr. Douglas Melton DNA sequencing by:
 Washington University Genome Sequencing Center For information on
 obtaining a clone please contact: Juliana Brown
 (brown@fas.harvard.edu) This sequence now available from the IMAGE
 consortium, for clone orders contact: info@image.llnl.gov
 Seq primer: -40UP from Gibco
 High quality sequence stop: 448.
 Location/Qualifiers
 1 .584

Search completed: February 25, 2004, 22:04:40
Job time : 4002 secs

```

ORIGIN
Library. "
Query Match 34.4%; SCORE 576; DB 12; Length 584;
Best Local Similarity 99.1%; Pred.No. 1.7e-84;
Matches 579; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1083 TGTGCTGTGAATCTACAGAAAGTCACAGAAATGCTTCTTAAAGGAGAAAGAGTTCCAC 1142

```